

**Strukturelle Charakterisierung zweier  
Schlüsselenzyme  
aus fermentativen Abbauwegen  
anaerober Bakterien**

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Chemie

der Philipps-Universität Marburg

vorgelegt von

Daniel Kreß

aus

Alzenau i. Ufr.

Marburg/Lahn 2009

Vom Fachbereich Chemie

der Philipps-Universität Marburg

als Dissertation angenommen am: \_\_\_\_\_

Erstgutachter: Prof. Dr. Lars-Oliver Essen (Fachbereich Chemie, Marburg)

Zweitgutachter: Prof. Dr. Wolfgang Buckel (Fachbereich Biologie, Marburg)

Tag der mündlichen Prüfung am: \_\_\_\_\_

# Erklärung

Ich versichere, dass ich meine Dissertation mit dem Thema:

**„Strukturelle Charakterisierung zweier Schlüsselenzyme  
aus fermentativen Abbauwegen anaerober Bakterien“**

selbständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als den von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

---

(Ort/Datum)

---

(Daniel Kreß)





**Teile dieser Arbeit sind Bestandteil der folgenden Publikationen:**

Kress D., Bruegel D., Buckel W., Essen L.-O. (2009). **An asymmetric model for Na<sup>+</sup>-translocating glutaconyl-CoA decarboxylases.** *Journal of Biological Chemistry*, im Druck, (online seit August 4, 2009; doi:10.1074/jbc.M109.037762).

Kress D., Alhapel A., Pierik A.J., Essen L.-O. (2008). **The crystal structure of enamidase: a bifunctional enzyme of the nicotinate catabolism.** *Journal of Molecular Biology*, 384, 837-47.

Kress D., Bozarth A., Pierik A.J., Essen L.-O., **Crystal structure of the *Bradyrhizobium japonicum*-enamidase.** *Manuskript in Vorbereitung.*

**Weitere Publikationen:**

Psakis G., Nitschkowski S., Holz C., Kress D., Maestre-Reyna M., Polaczek J., Illing G., Essen L.-O. (2007). **Expression screening of integral membrane proteins from *Helicobacter pylori* 26695.** *Protein Science*, 16 (12), 2667-76.

Schönafinger A., Morbitzer A., Kress D., Essen L.-O., Noll F., Hampp N. (2006). **Morphology of dry solid supported protein monolayers in dependence on substrate and protein surface properties.** *Langmuir* 22, 7185-91.

Kress D., Essen L.-O., **Crystal structure determination of the soluble domain of a component of the flagellar export apparatus by MAD-phasing.** *Manuskript in Vorbereitung.*

### Posterbeiträge (Auswahl):

**Kress D., Reitz S., Alhapel A., Pierik A.J., Essen L.-O. (2008). Structural characterization of key enzymes of the *Eubacterium barkeri* nicotinate fermentation pathway.** *Jahrestagung der Vereinigung für Allgemeine und Angewandte Mikrobiologie (VAAM) und der Gesellschaft für Biochemie und Molekular-biologie (GBM), Frankfurt a. Main.*

Psakis G., Nitschkowski S., Hölscher S., Neuhaus C., **Kress D.**, Maestre-Reyna M., Essen L.-O. (2007). **High-throughput expression and crystallisation of membrane proteins from *Helicobacter pylori* and homologues.** *Structural Proteomics of Membrane Proteins, Rauischholzhausen.*

## Abkürzungen und Fachbegriffe

Å	Ångström ( $1 \text{ Å} = 10^{-10} \text{ m}$ )
Amp	Ampicillin
APS	Ammoniumperoxodisulfat
ATP	Adenosin-5'-triphosphat
BSA	Rinderserumalbumin (engl.: <i>bovine serum albumine</i> )
CCD	engl.: <i>charge coupled device</i>
C-terminal	carboxy-terminal
Da	Dalton
DESY	Deutsches Elektronensynchrotron
DNA	Desoxyribonukleinsäure (engl.: <i>desoxyribonucleic acid</i> )
DTT	1,4-Dithiothreitol
EDTA	Ethylendiamintetraessigsäure
ESRF	engl.: <i>European Synchrotron Facility</i>
FPLC	engl.: <i>fast performance liquid chromatography</i>
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure
IP	engl.: <i>image plate</i>
K	Kelvin
LB	Luria-Bertani
M	molar (Mol/L)
MALDI-TOF	engl.: <i>matrix-assisted laser desorption/ionization-time of flight</i>
MES	2-(N-Morpholino)-ethansulfonsäure
MW	Molekulargewicht
MWCO	engl.: <i>molecular weight cut off</i>
NMR	magnetische Kernresonanz (engl.: <i>nuclear magnetic resonance</i> )
N-terminal	aminoterminal
OD <sub>x</sub>	optische Dichte bei der Wellenlänge x nm
PAGE	Polyacrylamidgelelektrophorese
PEG	Polyethylenglycol
PMSF	Phenylmethylen-sulfonylfluorid
r. m. s. d.	Standardabweichung (engl.: <i>root mean square deviation</i> )

rpm	Umdrehungen pro Minute (engl.: <i>revolutions per minute</i> )
RT	Raumtemperatur
SDS	Natriumdodecylsulfat (engl.: <i>sodium dodecyl sulfate</i> )
TEMED	N,N,N',N'-Tetramethylethyldiamin
Tris	Tris(hydroxymethyl)-aminomethan
v/v	Volumen pro Volumen
w/v	Gewicht pro Volumen (engl.: <i>weight per volume</i> )

## Aminosäuren: Abkürzungen und Molekulargewichte

Aminosäure	3-Buchstabencode	1-Buchstabencode	MW [g/mol]
Alanin	Ala	A	89
Arginin	Arg	R	174
Asparagin	Asn	N	132
Aspartat	Asp	D	133
Cystein	Cys	C	121
Glutamin	Gln	Q	146
Gyclin	Gly	G	75
Histidin	His	H	155
Isoleucin	Ile	I	131
Leucin	Leu	L	131
Lysin	Lys	K	146
Methionin	Met	M	149
Ornithin	Orn	O	132
Phenylalanin	Phe	F	165
Prolin	Pro	P	115
Serin	Ser	S	105
Threonin	Thr	T	119
Tryptophan	Trp	W	204
Tyrosin	Tyr	Y	181
Valin	Val	V	117

# Inhalt

<b>ZUSAMMENFASSUNG.....</b>	<b>1</b>
<b>SUMMARY .....</b>	<b>3</b>
<b>1 EINLEITUNG .....</b>	<b>5</b>
1.1 ENAMIDASE .....	6
1.1.1 <i>Nicotinat-Fermentation in Eubacterium barkeri</i> .....	6
1.1.2 <i>Enamidase, eine bifunktionelle Amidohydrolase</i> .....	8
1.1.3 <i>Amidohydrolasen</i> .....	10
1.2 GLUTACONYL-COENZYM-A-DECARBOXYLASE .....	15
1.2.1 <i>Glutamat-Abbau in anaeroben Bakterien</i> .....	15
1.2.2 <i>Die Rolle von Na<sup>+</sup> in der Kopplung bioenergetischer Prozesse</i> .....	19
1.2.3 <i>Die Natriumionen-Pumpe Glutaconyl-CoA-Decarboxylase</i> .....	21
1.3 ZIELSETZUNG.....	27
<b>2 MATERIAL .....</b>	<b>29</b>
2.1 GERÄTE, VERBRAUCHSMATERIALIEN UND CHEMIKALIEN .....	29
2.2 VEKTOREN UND MIKROORGANISMEN.....	33
2.3 NÄHRMEDIE, STAMMLÖSUNGEN UND PUFFER.....	35
<b>3 METHODEN .....</b>	<b>39</b>
3.1 MOLEKULARBIOLOGISCHE METHODEN .....	39
3.2 PROTEINCHEMISCHE METHODEN.....	39
3.2.1 <i>Heterologe Genexpression in E. coli</i> .....	40
3.2.2 <i>Zellaufschluss</i> .....	40
3.2.3 <i>Chromatographische Proteinaufreinigung</i> .....	41
3.2.4 <i>Konzentration der Proteinproben</i> .....	42
3.2.5 <i>Entsalzung und Umpufferung</i> .....	42
3.2.6 <i>Proteinkonzentrationsbestimmung</i> .....	42
3.2.7 <i>SDS-Polyacrylamidgelelektrophorese (SDS-PAGE)</i> .....	43
3.3 PROTEINKRISTALLISATION .....	45
3.3.1 <i>Automatisiertes Kristallisationsscreening</i> .....	47
3.3.2 <i>Kristallisationsoptimierung</i> .....	48
3.3.3 <i>Kokristallisation und soaking-Experimente</i> .....	48

3.4	PROTEINSTRUKTURBESTIMMUNG.....	51
3.4.1	Kristalle und Symmetrie .....	51
3.4.2	Erzeugung von Röntgenstrahlen .....	55
3.4.3	Beugung von Röntgenstrahlen an Kristallen.....	57
3.4.4	Reziprokes Gitter und Ewald-Konstruktion.....	59
3.4.5	Detektion gebeugter Röntgenstrahlung .....	62
3.4.6	Strukturfaktorgleichung, Elektronendichte- und Patterson-Funktion .....	63
3.4.7	Phasenbestimmung durch Molekularen Ersatz .....	66
3.4.8	Phasenbestimmung durch multiple anomale Dispersion.....	72
3.4.9	Datensammlung.....	77
3.4.10	Skalierung und Reduktion der Daten .....	80
3.4.11	Dichtemodifikation .....	83
3.4.12	Modellbau.....	86
3.4.13	Verfeinerung des Strukturmodells .....	89
3.4.14	Bioinformatische Charakterisierung der Strukturmodelle .....	91
4	ERGEBNISSE .....	95
4.1	DIE KRISTALLSTRUKTUR DER EUBACTERIUM BARKERI-ENAMIDASE.....	95
4.1.1	Klonierung, Expression und Aufreinigung .....	95
4.1.2	Kristallisation.....	96
4.1.3	Datensammlung und -auswertung.....	97
4.1.4	Strukturlösung und -verfeinerung .....	98
4.1.5	Qualität des Strukturmodells.....	101
4.1.6	Beschreibung des Enamidase-Strukturmodells .....	101
4.1.7	Quartärstruktur .....	103
4.1.8	Das aktive Zentrum der Enamidase.....	104
4.2	DIE KRISTALLSTRUKTUR DER BRADYRHIZOBIUM JAPONICUM-ENAMIDASE.....	109
4.2.1	Klonierung, Expression und Aufreinigung .....	109
4.2.2	Kristallisation, Datensammlung und Strukturlösung.....	110
4.3	DIE KRISTALLSTRUKTUR DER CLOSTRIDIUM SYMBIOSUM-GCDA .....	113
4.3.1	Klonierung, Expression und Aufreinigung .....	113
4.3.2	Kristallisation.....	113
4.3.3	Datensammlung und -auswertung.....	115
4.3.4	Strukturlösung und -verfeinerung .....	116
4.3.5	Qualität der Strukturmodelle.....	117
4.3.6	Beschreibung des Strukturmodells für das GcdA-Monomer .....	119

4.3.7	<i>Architektur des aktiven Zentrums der GcdA</i> .....	122
4.3.8	<i>Die Glutaconyl-CoA-Bindungstasche</i> .....	123
4.3.9	<i>Quartärstruktur</i> .....	125
<b>5</b>	<b>DISKUSSION</b> .....	<b>127</b>
5.1	ENAMIDASE .....	127
5.1.1	<i>Vergleich der Enamidase mit anderen Amidohydrolasen</i> .....	127
5.1.2	<i>Substraterkennung</i> .....	130
5.1.3	<i>Der Mechanismus der Enamidase-Reaktion</i> .....	132
5.1.4	<i>Ausblick</i> .....	133
5.2	GLUTACONYL-CoA-DECARBOXYLASE .....	135
5.2.1	<i>Mechanismus der Decarboxylase-Untereinheit</i> .....	135
5.2.2	<i>Ein asymmetrisches Modell für Na<sup>+</sup>-pumpende Glutaconyl-CoA-Decarboxylasen</i> .....	136
5.2.3	<i>Ausblick</i> .....	140
<b>6</b>	<b>LITERATURVERZEICHNIS</b> .....	<b>141</b>
<b>7</b>	<b>DANKSAGUNG</b> .....	<b>157</b>





## Zusammenfassung

### Enamidase

Die Hydrolyse von 1,4,5,6-Tetrahydro-6-oxonicotinat (THON) zu 2-Formylglutarat ist ein Schlüsselschritt im Nicotin-Katabolismus einer Reihe von Clostridien und Proteobakterien, der von dem Enzym Enamidase (EC 3.5.2.18) katalysiert wird. Auf Grund von Sequenzvergleichen, der Art der katalysierten Reaktion und dem Nachweis stöchiometrischer Bindung von Metallionen wurde es der Amidohydrolase-Strukturfamilie zugeordnet. Enamidase katalysiert eine ungewöhnliche Zweistufenreaktion: die Dezyklisierung von THON zu 2-(Enamin)-glutarat und dessen weitere Hydrolyse zu (S)-2-Formylglutarat. Die hier präsentierte 1.9 Å-Kristallstruktur der Enamidase aus *Eubacterium barkeri* liefert die strukturelle Grundlage für den Katalysemechanismus dieser enantioselektiven Reaktion. Das Enzym bildet 222-symmetrische Tetramere. Die Enamidase-Monomere bestehen dabei aus einer  $\beta$ -Sandwich-Domäne, die sich aus den N- und C-Termini der Polypeptidkette zusammensetzt, und einer zentralen  $(\alpha/\beta)_8$ -Fass-Domäne die das aktive Zentrum enthält. Letzteres beinhaltet ein binukleares Metallzentrum aus Zink- und Eisenionen. Damit stellt Enamidase eine spezielle Typ-II-Amidohydrolase dar.

### Decarboxylase-Untereinheit der Natriumionen-Pumpe Glutaconyl-CoA-Decarboxylase

Glutaconyl-CoA-Decarboxylase (Gcd, EC 4.1.1.70) koppelt die Decarboxylierung von Glutaconyl-CoA mit dem Aufbau eines  $\text{Na}^+$ -Gradienten. Das integrale Membranenzym besteht aus vier Untereinheiten: der  $\alpha$ -Untereinheit (GcdA), die die Carboxyl-Gruppe von Glutaconyl-CoA auf die biotinylierte  $\gamma$ -Untereinheit (GcdC) transferiert, der  $\beta$ -Untereinheit (GcdB), die die Decarboxylierung des Carboxybiotins sowie die  $\text{Na}^+$ -Translokation katalysiert, und der  $\delta$ -Untereinheit (GcdD), deren Funktion noch nicht bekannt ist. Die in der vorliegenden Arbeit präsentierten Kokristallstrukturen der GcdA-Untereinheit aus *Clostridium symbiosum* mit dem Substrat Glutaconyl-CoA, dem Produkt Crotonyl-CoA und dem Substratanalogon Glutaryl-CoA liefern ein hochaufgelöstes Modell für die Quartär-struktur und den Katalysemechanismus des Enzyms, das bemerkenswerte strukturelle Veränderungen bei der Substratbindung aufdeckt. Auf der Grundlage dieser Daten wird ein neuartiges, asymmetrisches Modell für den intakten Gcd-Komplex vorgeschlagen, in dem GcdA ein Tetramer bildet, das von einem Netzwerk Lösungsmittel-gefüllter Kanäle durchzogen ist.



## Summary

### Enamidase

The hydrolysis of 1,4,5,6-tetrahydro-6-oxonicotinate to 2-formylglutarate is a central step in the catabolism of nicotinate in several Clostridia and Proteobacteria. This reaction is catalyzed by the novel enzyme enamidase (EC 3.5.2.18), a new member of the amidohydrolase superfamily as indicated by its unique reaction, sequence comparison, and the stoichiometric binding of metal ions. A hallmark of enamidase is its capability to catalyze a two-step reaction: the initial decyclization of 1,4,5,6-tetrahydro-6-oxonicotinate leading to 2-(enamine)glutarate followed by an additional hydrolysis step yielding (S)-2-formylglutarate. This study presents the crystal structure of enamidase from *Eubacterium barkeri* at 1.9 Å, providing a structural basis for catalysis and suggesting a mechanism for its exceptional activity and enantioselectivity. The enzyme forms a 222-symmetric tetramer built up by a dimer of dimers. Each enamidase monomer consists of a composite beta-sandwich domain comprising the N- and C-termini of the polypeptide chain and an  $(\alpha/\beta)_8$ -TIM-barrel domain harboring the active site. With its catalytic binuclear metal center comprising both zinc and iron ions, enamidase represents a special case of subtype II amidohydrolases.

### Decarboxylase subunit of the sodium ion pump glutaconyl-CoA decarboxylase

Glutaconyl-CoA decarboxylase (Gcd, EC 4.1.1.70) couples the biotin-dependent decarboxylation of glutaconyl-CoA with the generation of an electrochemical Na<sup>+</sup> gradient. As an integral membrane enzyme Gcd comprises four different subunits: the  $\alpha$ -subunit (GcdA) catalyzing the transfer of the carboxylate of glutaconyl-CoA to the biotinylated  $\gamma$ -subunit (GcdC), the membrane resident  $\beta$ -subunit (GcdB) responsible for the subsequent decarboxylation of carboxybiotin coupled to Na<sup>+</sup> translocation and a small  $\delta$ -subunit (GcdD) of yet unknown function. Co-crystallization of the decarboxylase subunit GcdA with the substrate glutaconyl-CoA, the product crotonyl-CoA and the substrate analogue glutaryl-CoA, respectively, resulted in a high resolution model for quaternary structure and catalysis of the enzyme revealing remarkable structural changes upon substrate binding. On the basis of these data a new asymmetric model for the intact Gcd complex is presented in which GcdA is associated as a tetramer crisscrossed by a network of solvent filled channel.



# 1 Einleitung

Alle Lebewesen benötigen Energie für die in ihren Zellen ablaufenden Prozesse. Die universelle „Währung“ des intrazellulären Energiehaushalts aller Organismen stellt das Molekül Adenosintriphosphat (ATP) dar [Knowles 1980]. Die meisten Eukaryoten produzieren ATP hauptsächlich mit Hilfe eines der folgenden Prozesse: Alkohol-Fermentation (z.B. Hefen), Lactat-Fermentation (z.B. Muskelzellen), aerobe Atmung (z.B. Schimmelpilze, Tiere) oder oxygene Photosynthese (z.B. Algen, Pflanzen). Jeder dieser Prozesse kommt auch in Prokaryoten vor. Allerdings verfügen diese darüber hinaus auch noch über zahlreiche alternative Arten der Energiegewinnung, die in eukaryotischen Zellen praktisch nicht vorkommen. Dazu zählen unter anderem die anaerobe Atmung, die anoxygene Photosynthese, die Lithotrophie, die Methanogenese sowie eine große Anzahl einzigartiger Fermentationswege.

Als Fermentationen werden in biologischen Systemen Prozesse bezeichnet, bei denen Energie durch partielle Oxidation organischer Moleküle, wie beispielsweise Kohlenhydrate oder Aminosäuren, gewonnen wird. Dabei fungieren Reaktionsintermediate des Stoffwechselweges als endogene Elektronenakzeptoren und -donoren, und ATP wird im Gegensatz zur „normalen“ zellulären Atmung nicht durch die oxidative Phosphorylierung im Rahmen der Atmungskette, sondern durch Phosphatkettenphosphorylierung erzeugt.

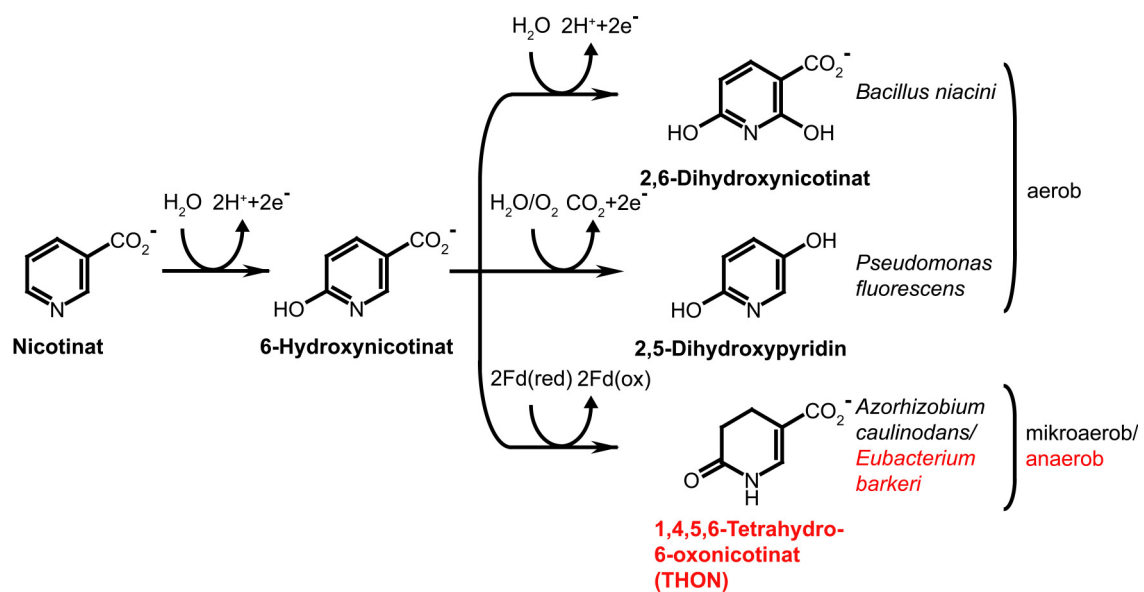
Bakterien sind in der Lage, aus unterschiedlichsten Substraten Energie durch Fermentation zu erzeugen. Dabei spielt eine Fülle von speziellen Zwischenprodukten und außergewöhnlichen Enzymen eine Rolle, die nicht nur dem Mikrobiologen und Strukturbiochemiker ein interessantes Betätigungsfeld bieten, sondern auch für die industrielle Synthese organischer Verbindungen im Rahmen der *weißen Biotechnologie* von zunehmendem Interesse sind [Buckel 2001b; Michael & Rainer 2006]. Ein Paradebeispiel für die Vielseitigkeit des mikrobiellen Katabolismus stellt der Abbau von Aminosäuren durch fermentative Bakterien dar: Für alle 20 proteinogenen Aminosäuren ist mindestens ein katabolischer Stoffwechselweg bekannt [Barker 1961; Barker 1981].

Gegenstand der vorliegenden Arbeit ist die strukturelle Charakterisierung zweier Schlüsselenzyme fermentativer Abbauwege in anaeroben Bakterien: des bifunktionellen Enzyms Enamidase aus der Nicotinat-Fermentation in *Eubacterium barkeri* und der Decarboxylase-Untereinheit des Enzymkomplexes Glutaconyl-CoA-Decarboxylase aus der Glutamat-Fermentation in *Clostridium symbiosum*.

## 1.1 Enamidase

### 1.1.1 Nicotinat-Fermentation in *Eubacterium barkeri*

Als Baustein von Nicotinsäureamid-Adenin-Dinukleotid-Kofaktoren, wie  $\text{NAD}^+$  und  $\text{NADP}^+$ , ist Nicotinsäure (Nicotinat, Niacin, Vitamin B3) in der Natur allgegenwärtig. Durch das Absterben von Organismen und den anschließenden Zerfall der NAD-Kofaktoren steht Nicotinat entsprechend spezialisierten Mikroorganismen als Stickstoff-, Kohlenstoff- und Energiequelle zur Verfügung. Der mikrobielle Abbau beginnt dabei stets mit der Oxidation des Nicotinats zu 6-Hydroxynicotinat (vgl. Abbildung 1-1). Diese Hydroxylierung wird von dem Enzym Nicotinat-Dehydrogenase katalysiert. Der weitere Abbau des Nicotinats richtet sich nach dem Sauerstoffangebot in der natürlichen Umgebung der jeweiligen Mikroorganismen.



**Abbildung 1-1 Nicotinat-Katabolismus in Mikroorganismen** Alle bekannten Abbauege beginnen mit der durch die Nicotinat-Dehydrogenase katalysierten Bildung von 6-Hydroxynicotinat. Die weitere Degradation erfolgt je nach der Verfügbarkeit von Sauerstoff über unterschiedliche Wege.

Bakterien der Spezies *Bacillus* und *Pseudomonas* sind in der Lage, Nicotinat unter aeroben Bedingungen abzubauen [Behrman & Stanier 1957; Ensign & Rittenberg 1964]. Dieser Prozess wird in der industriellen Biosynthese von 6-Hydroxynicotinat durch *Alcaligenes* sp. großtechnisch angewendet [Kulla 1991]. Die anaerobe Degradation ist deutlich weniger weit

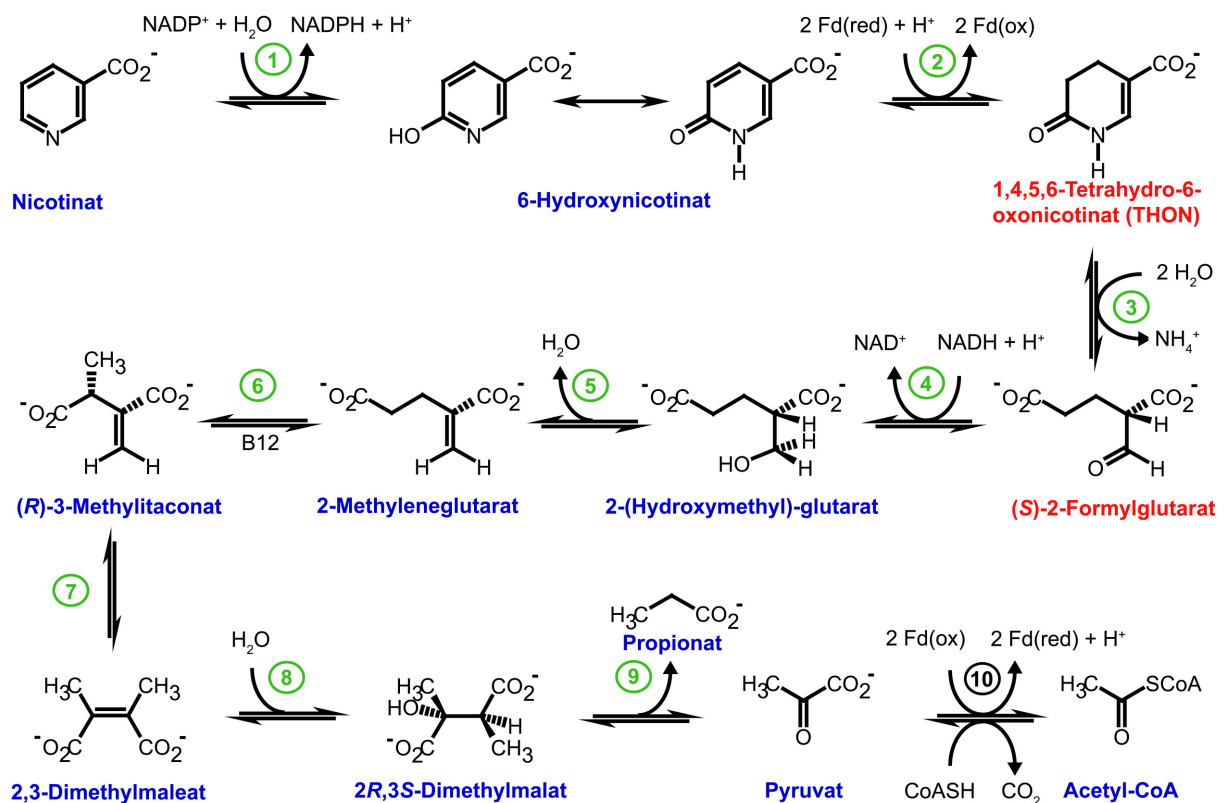
verbreitet. So ist bisher nur für einige Bakterien, wie beispielsweise *Azorhizobium caulinodans* [Kitts *et al.* 1992] und *Desulfobacterium niacini* [Imhoff-Stuckle & Pfennig 1983], nachgewiesen worden, dass sie Nicotinat unter mikroaeroben bzw. Sulfat-reduzierenden Bedingungen verstoffwechseln können.

Die umfassendsten Arbeiten auf dem Gebiet des anaeroben Nicotinat-Abbaus wurden für *Eubacterium barkeri* durchgeführt. Dieses anaerobe Bodenbakterium wurde 1957 erstmals aus Sedimenten des Flusses Potomac in Washington, D.C. (USA) isoliert [Harary 1957] und gehört zur physiologischen Gruppe der saccharolytischen *Clostridien* [Collins *et al.* 1994; Stadtman *et al.* 1972]. Der Stoffwechselweg der Nicotinat-Fermentation in *E. barkeri* wurde von Kung *et al.* [Kung *et al.* 1971] und Eggerer [Eggerer 1985] bereits teilweise aufgeklärt und weist folgende Bilanz auf:



Wie bei vielen anderen Fermentationen ist die Energieausbeute dabei gering: *Eubacterium barkeri* gewinnt lediglich ein einziges ATP pro Nicotinat-Molekül durch Substratkettenphosphorylierung. Eine detaillierte molekulare Charakterisierung dieses einzigartigen Fermentationsweges präsentierten Alhapel *et al.* [Alhapel *et al.* 2006]. In dieser Studie wurden der Gencluster, der die Gene aller neun beteiligten Enzyme enthält, kloniert und sequenziert und neben einer Reihe von Reaktionszwischenstufen drei neue Enzyme identifiziert: 6-Hydroxynicotinat-Reduktase, 2-(Hydroxymethyl)-glutarat-Dehydrogenase und Enamidase. Der vollständige Fermentationsweg mit sämtlichen Zwischenstufen und allen beteiligten Enzymen ist in Abbildung 1-2 dargestellt.

Enamidase katalysiert die Dezyklisierung und enantioselektive Hydrolyse von 1,4,5,6-Tetrahydro-6-oxonicotinate (THON), dem Produkt der 6-Hydroxynicotinat-Reduktase, zu (S)-2-Formylglutarat. Die strukturelle Charakterisierung dieses außergewöhnlichen Enzyms der Amidohydrolase-Familie ist Gegenstand der hier vorgestellten Arbeit.



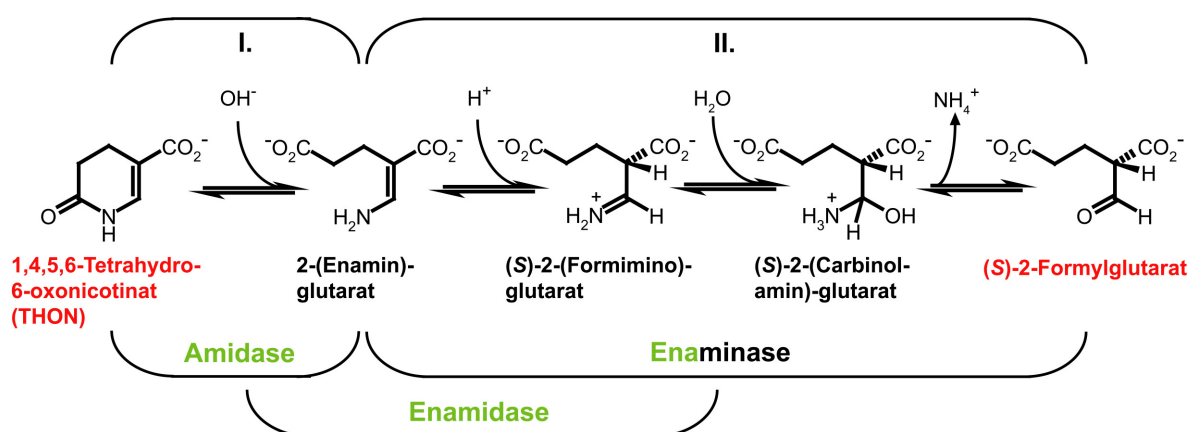
**Abbildung 1-2 Anaerober Nicotinat-Abbau in *Eubacterium barkeri*** Die grünen Zahlen kennzeichnen die Enzyme des Nicotinat-Katabolismus-Genclusters: **1**, Nicotinat-Dehydrogenase (EC 1.17.1.5); **2**, 6-Hydroxynicotinat-Reductase (EC 1.3.7.1); **3**, Enamidase (EC 3.5.2.18); **4**, 2-(Hydroxymethyl)-glutarat-Dehydrogenase (EC 1.1.1.291); **5**, 2-(Hydroxymethyl)-glutarat-Dehydratase; **6**, 2-Methyleneglutarat-Mutase (EC 5.4.99.4); **7**, Methylitaconat-Δ-Isomerase (EC 5.3.3.6); **8**, Dimethylmaleat-Hydratase (EC 4.2.1.85); **9**, 2,3-Dimethylmalat-Lyase (EC 4.1.3.32). Abschließend wird das gebildete Pyruvat-Molekül von der Pyruvat-Ferredoxin-Oxidoreductase (**10**, Pyruvat-Synthase, EC 1.2.7.1) zu Acetyl-CoA umgesetzt, welches durch Substratkettenphosphorylierung ein Molekül ATP liefert.

## 1.1.2 Enamidase, eine bifunktionelle Amidohydrolase

Enamidase katalysiert die Hydrolyse von THON zu (S)-2-Formylglutarat und damit die dritte Reaktion der Nicotinat-Fermentation in *E. barkeri*. Bei dieser Hydrolyse handelt es sich um eine Zweistufenreaktion (Abbildung 1-3): Die Dezyklisierung von THON führt zunächst zur Bildung von 2-(Enamin)-glutarat, welches anschließend zu (S)-2-Formylglutarat weiterhydrolysiert wird. Der Name Enamidase rührt daher, dass sich der erste Teil der Reaktion formal als *Amidase*-Funktionalität beschreiben lässt, während der zweite Teil einer *Enaminase*-Aktivität entspricht. Alhapel et al. konnten sowohl für das aus *E. barkeri* aufgereinigte als auch für das in *E. coli* rekombinant überproduzierte Enzym zeigen, dass der zweite Teil der



Reaktion, die Umsetzung des Hydrolyse-empfindlichen Enamins, enantioselektiv und damit tatsächlich enzymkatalysiert verläuft. So wurde nachgewiesen, dass das NADH-abhängige Folgeenzym, die 2-(Hydroxymethyl)-glutarat-Dehydrogenase, ausschließlich (S)-2-Formylglutarat als Substrat für die Umsetzung zu (S)-2-(Hydroxymethyl)-glutarat akzeptiert. Die Stereospezifität dieser Reaktion wurde verwendet, um die Chiralität des von Enamidase gebildeten (S)-2-Formylglutarats zu bestimmen.



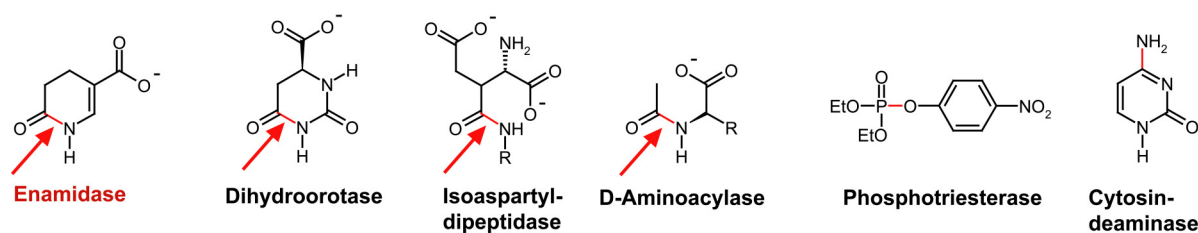
**Abbildung 1-3 Die von Enamidase katalysierte Zweistufenreaktion (I.)** Im ersten Schritt der Reaktion erfolgt die Ring-öffnende Hydrolyse von THON zu 2-(Enamin)-glutarat. (II.) Die sich anschließende Hydrolyse vollzieht sich über die Reaktionsintermediate (S)-2-(Formimino)-glutarat und (S)-2-(Carbinolamin)-glutarat und liefert ausschließlich (S)-2-Formylglutarat. Der Name *Enamidase* leitet sich von der formalen Unterteilung der Enzym-Funktionalität in *Amidase*- und *Enaminase*-Katalyse ab.

Die aus *E. barkeri* aufgereinigte Enamidase hat ein berechnetes Molekulargewicht von 39,793 Da und eluiert in der Volumenausschlusschromatographie bei einem apparenten Molekulargewicht, das einem Homotetramer entspricht. Für die Hydrolyse von THON wurde ein  $K_m$  von 5 mM ermittelt. Darüber hinaus konnte durch colorimetrische Experimente gezeigt werden, dass die aus *E. barkeri* aufgereinigte Enamidase 1.0 Äquivalente Eisen und ca. 0.6 Äquivalente Zink enthält. Sequenzvergleiche, der Nachweis annähernd stöchiometrischer Mengen Eisen und Zink sowie die Art der katalysierten Reaktion legten die Schlussfolgerung nahe, dass es sich bei Enamidase wahrscheinlich um einen neuen Vertreter der Strukturfamilie der Amidohydrolasen handelt [Alhapel *et al.* 2006].

### 1.1.3 Amidohydrolasen

Die Strukturfamilie der Amidohydrolasen wurde 1997 von Holm und Sander eingeführt [Holm & Sander 1997]. Auf Grund der auffälligen Ähnlichkeiten zwischen den Strukturmodellen für Adenosin-Deaminase, Phosphotriesterase und Urease fassten sie eine große Anzahl von  $d$ -Block-Metall-abhängigen hydrolytischen Enzymen zu dieser neuen Strukturfamilie zusammen.

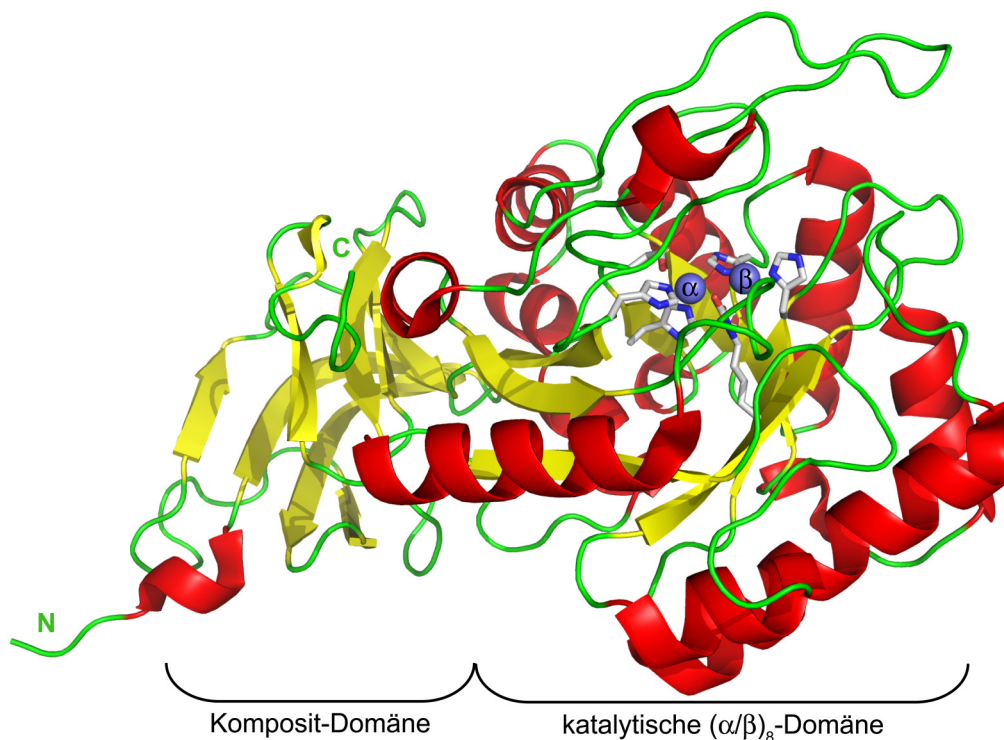
Die Amidohydrolasen werden in der strukturbasierten SCOP-Datenbank (Structural Characterization of Proteins, <http://scop.mrc-lmb.cam.ac.uk/scop/>, [Murzin *et al.* 1995]) als Metall-abhängige Hydrolasen geführt und bilden eine große Untergruppe der „Alpha- und Beta-Proteine ( $\alpha/\beta$ )“. In der Pfam-Datenbank (Protein families, <http://pfam.sanger.ac.uk/>, [Finn *et al.* 2008]) werden sie in drei Familien unterteilt. Die erste Familie umfasst momentan 6089 (Stand 9/2009) verschiedene Sequenzen und beinhaltet damit alle bekannten „klassischen“ Vertreter der Enzymfamilie. Auch die im Rahmen dieser Arbeit strukturell charakterisierte Enamidase gehört zu diesem Typ. Neben einer großen Bandbreite von Reaktionen, die die Hydrolyse von Amidbindungen betreffen, katalysieren die Enzyme dieser Gruppe auch die Desaminierung aromatischer Basen und die Hydrolyse von Nicht-Amidbindungen, wie beispielsweise Phosphor-Sauerstoff-Bindungen organischer Phosphorsäuretriester-Derivate (Abbildung 1-4). In der zweiten Gruppe sind 1497 Sequenzen enthalten, die mit der ersten Familie verwandt sind, sich aber strukturell oder funktionell von ihr unterscheiden. Vertreter dieser Gruppe katalysieren beispielsweise nicht-hydrolytische Reaktionen, wie Decarboxylierungen und Hydrierungen. In der dritten Pfam-Familie werden insgesamt 1432 Sequenzen zusammengefasst. Bei dieser Gruppe handelt es sich größtenteils um hypothetische Proteine, die eine gewisse Verwandtschaft zu den Amidohydrolasen aufweisen, jedoch nur über eine geringe Sequenzidentität mit diesen verfügen.



**Abbildung 1-4 Von Enamidase und ausgewählten Amidohydrolasen katalysierte Reaktionen**  
Die von den Enzymen gespaltenen Bindungen sind jeweils rot hervorgehoben.

## Struktur

Die allgemeine Struktur von Amidohydrolasen ist in Abbildung 1-5 anhand der Monomerstruktur eines typischen Vertreters dieser Familie, der Isoaspartyl-Dipeptidase (PDB-Code: 1poj), beispielhaft illustriert. Alle Amidohydrolasen zeigen als charakteristisches Strukturmotiv ein mehr oder weniger stark verzerrtes  $(\alpha/\beta)_8$ -Fass-Motiv. Dieses Faltungsmotiv ist bei ca. 10% aller strukturell charakterisierten Proteine und in ca. 30 Strukturfamilien der SCOP-Datenbank zu finden und wird in Anlehnung an Triosephosphat-Isomerase oft auch als *TIM-barrel* bezeichnet [Aimin *et al.* 2007]. Darüber hinaus beinhalten die Strukturen einiger Amidohydrolasen eine zweite Domäne, die von  $\beta$ -Strängen dominiert wird und häufig in Form eines  $\beta$ -Sandwichs gefaltet ist. Für die Katalyse ist allerdings ausschließlich die  $(\alpha/\beta)_8$ -Domäne von Bedeutung. Sie trägt ein mono- bzw. binukleares Metallzentrum, an dem die jeweilige Reaktion katalysiert wird.



**Abbildung 1-5 Monomerstruktur der Isoaspartyl-Dipeptidase (PDB-Code: 1poj)**

Die Isoaspartyl-Dipeptidase gehört zu den Typ I-Amidohydrolasen und zeigt die für viele Amidohydrolasen typische Zweidomänenstruktur. Die  $\beta$ -Sandwich-Domäne im linken Teil der Struktur wird auch als Kompositdomäne bezeichnet, da sie aus N- und C-terminalen Enden der Polypeptidkette gebildet wird. Die katalytische  $(\alpha/\beta)_8$ -Fass-Domäne beinhaltet das binukleare Metallzentrum mit zwei Zinkatomen an den Positionen  $\alpha$  und  $\beta$ . Die Schleifenregionen,  $\alpha$ -Helices und  $\beta$ -Stränge sind grün, rot bzw. gelb hervorgehoben. Die Reste des binuklearen Metallzentrums sind in grau und die Zinkionen als blaue Kugeln dargestellt.

## Metallzentrum

Das Metallzentrum befindet sich an einem Ende der  $(\alpha/\beta)_8$ -Fass-Struktur in der Nähe der C-Termini der  $\beta$ -Stränge. Die Seitenketten von bis zu sechs Aminosäureresten koordinieren ein oder zwei zweiwertige Metallionen der  $d$ -Reihe. Neben Zink werden dabei auch Eisen- und Nickelionen beobachtet. Entsprechend der Anzahl und Art der metallbindenden Reste sowie ihrer Zugehörigkeit zu den jeweiligen  $\beta$ -Strängen können sieben Typen von Amidohydrolasen unterschieden werden (vgl. Tabelle 1-1) [Seibert & Raushel 2005].

Typ	Metall	Position	$\beta$ -Strang							
			1	2	3	4	5	6	7	8
I	Zn, Ni	$\alpha, \beta$	HXH			K <sup>c</sup>	H	H		D
II	Zn	$\alpha, \beta$	HXH			E	H	H		D
III	Zn, Fe	$\alpha$	HXH				H	h <sup>b</sup>		D
IV	Fe	$\beta$	hxh <sup>a</sup>		E		H	H		d <sup>b</sup>
V	Zn	$\beta$	hxh <sup>a</sup>	C			H	H		d <sup>b</sup>
VI	Zn	$\alpha, \beta$	HXD		E		H	H		d <sup>b</sup>
VII			HXH				H			D

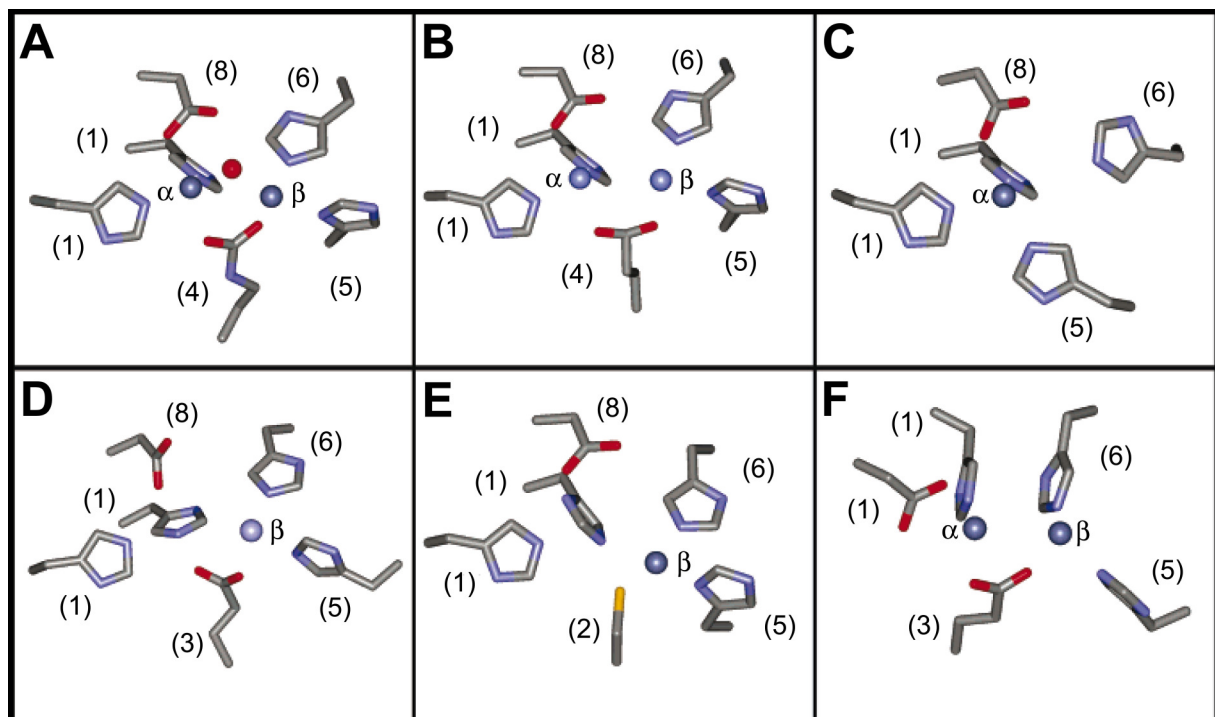
**Tabelle 1-1 Die sieben Typen von Metallliganden der mono- bzw. binuklearen Metallzentren innerhalb der Amidohydrolase-Familie [Seibert & Raushel 2005]**

<sup>a</sup> Diese Reste befinden sich im aktiven Zentrum, binden jedoch nicht direkt an das Metallion.

<sup>b</sup> Diese Reste binden nicht direkt an eines der Metallionen, bilden jedoch Wasserstoffbrückenbindungen zum hydrolytischen Wassermolekül aus.

<sup>c</sup> Es handelt sich hierbei stets um ein posttranslational carboxyliertes Lysin.

Das Metallzentrum von TypI-Amidohydrolasen ist in Abbildung 1-6A am Beispiel der Phosphotriesterase gezeigt: Die beiden Zinkatome in den Positionen  $\alpha$  und  $\beta$  werden von insgesamt vier Histidinen, einem Aspartat und der Carbamat-Seitenkette eines posttranslational carboxylierten Lysins koordiniert. Letzteres agiert dabei als verbrückender Ligand zwischen den Metallionen. Die übrigen Koordinationstypen ergeben sich im Wesentlichen durch die Variation des verbrückenden Aminosäurerests (vgl. Abbildung 1-6B-F). Am häufigsten handelt es sich dabei um ein Glutamat der  $\beta$ -Stränge 3 oder 4. Im mononuklearen Metallzentrum der D-Aminoacylase (Typ V) hingegen befindet sich an dieser Stelle ein Cystein-Rest des zweiten  $\beta$ -Strangs. Bei Adenosindeaminase (Typ III) fehlt dieser Rest gänzlich, und es liegen nur fünf Seitenketten in unmittelbarer Nähe zum einzelnen Metallion in Position  $\alpha$ .



**Abbildung 1-6 Sechs Typen mono- bzw. binuklearer Metallzentren ( $\alpha$ ,  $\beta$ ) in Amidohydrolasen**  
Gezeigt ist jeweils ein repräsentatives Beispiel: (A) Phosphotriesterase (die rote Kugel entspricht einem Hydroxidion) (B) Phosphotriesterase-ähnliches Protein, (C) Adenosin-Deaminase, (D) N-Acetylglucosamin-6-phosphat-Deacetylase, (E) D-Aminoacylase und (F) Dipeptidase aus der menschlichen Niere. Die Nummern beziehen sich auf die  $\beta$ -Stränge, von denen die jeweiligen Metallliganden stammen. Abbildung nach [Seibert & Raushel 2005].

## Reaktionsmechanismen

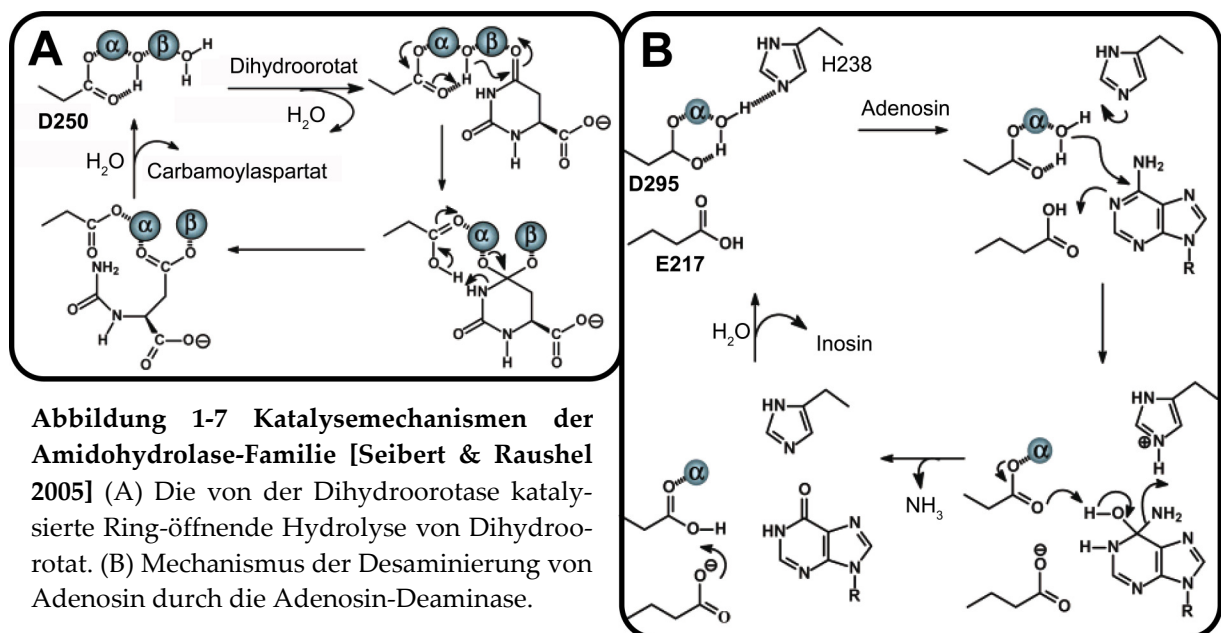
Zu den am besten untersuchten Reaktionsmechanismen von Amidohydrolasen zählen die von der *E. coli*-Dihydroorotase und der *Mus musculus*-Adenosin-Deaminase katalysierten hydrolytischen Reaktionen, die in Abbildung 1-7 skizziert sind. Die Dihydroorotase-Reaktion entspricht der Hydrolyse einer Amidbindung an einem binuklearen Metallzentrum. Aus einer Kristallstruktur, die sowohl Dihydroorotat als auch das Hydrolyse-Produkt Carbamoylaspartat enthielt, wurde der im Folgenden beschriebene Mechanismus abgeleitet (Abbildung 1-7A) [Thoden *et al.* 2001].

Die Reaktion beginnt mit der Koordination des Carbonyl-Sauerstoffatoms der zu spaltenden Amidbindung an das  $\beta$ -Zink des Metallzentrums. In vielen Kristallstrukturen von Amidohydrolasen mit binuklearem Metallzentrum konnte zwischen den Metallionen ein verbrückendes Wassermolekül nachgewiesen werden. Da davon ausgegangen wird, dass dieses als Hydroxidion vorliegt, kann es in der hydrolytischen Reaktion als Nukleophil zur Verfügung

stehen. In Dihydroorotase interagiert dieses Hydroxidion zusätzlich über eine Wasserstoffbrückenbindung mit dem Metallliganden Asp250, welches gewissermaßen als „Wasserstoff-Shuttle“ für die folgende Reaktion fungiert: Nach dem nukleophilen Angriff des Hydroxids am Kohlenstoffatom der polarisierten Carbonylgruppe wird der Amid-Stickstoff von Asp250 protoniert, wobei gleichzeitig die Hydrolyse der Bindung erfolgt.

Die Desaminierung von Adenosin durch die Adenosin-Deaminase ist ein Beispiel für die Katalyse in Gegenwart von nur einem Metallion im reaktiven Zentrum (Abbildung 1-7B) [Wang & Quijcho 1998; Wilson *et al.* 1991]. Hier wird zunächst einem an das  $\alpha$ -Metallion gebundenen Wassermolekül durch His238 ein Wasserstoffatom entzogen. Die gleichzeitige Protonierung des N1-Stickstoffatoms des Adenosins durch die Seitenkette des Glu217 begünstigt anschließend den nukleophilen Angriff des entstandenen Hydroxidions. Schließlich wird die Amin-Abgangsgruppe von His238 protoniert und leitet damit den Zerfall des tetraedrischen Intermediats und somit die Bildung von Inosin ein.

Beide beschriebenen Reaktionsmechanismen sind für die zu untersuchende Enamidase-Reaktion von Bedeutung: Dihydroorotat hat eine ähnliche Struktur wie das Substrat THON, und seine Hydrolyse gleicht dem ersten Teil der von Enamidase katalysierten Reaktion; die Desaminierung von Adenosin hingegen ist vergleichbar mit der Hydrolyse des Enamins im zweiten Schritt der Enamidase-Reaktion.



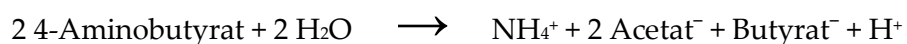
**Abbildung 1-7 Katalysemechanismen der Amidohydrolase-Familie [Seibert & Raushel 2005]** (A) Die von der Dihydroorotase katalysierte Ring-öffnende Hydrolyse von Dihydroorotat. (B) Mechanismus der Desaminierung von Adenosin durch die Adenosin-Deaminase.

## 1.2 Glutaconyl-Coenzym-A-Decarboxylase

### 1.2.1 Glutamat-Abbau in anaeroben Bakterien

In Eukaryoten erfolgt der Abbau von Glutamat stets über das Zwischenprodukt 2-Oxoglutarat ( $\alpha$ -Ketoglutarat), das anschließend durch die Enzyme des Krebs-Zyklus weiter oxidiert wird. Auch aerobe und anaerob atmende Bakterien nutzen diesen Stoffwechselweg. In fermentativen Bakterien der Ordnungen Clostridiales und Fusobacteriales dagegen sind bislang schon fünf alternative Wege bekannt, um aus Glutamat Energie zu gewinnen [Boiangiu *et al.* 2005; Buckel 2001a; Buckel 2001b]. Im Einzelnen unterscheidet man den 2-Hydroxyglutarat-, den 4-Aminobutyrat-, den 5-Aminovaleriat- sowie die Butyrat- und die Propionat-Variante des 3-Methylaspartat Wegs. Sie sind jeweils nach charakteristischen (Zwischen-)Produkten benannt.

Die Reaktionen des 2-Hydroxyglutarat- und der beiden Koenzym B<sub>12</sub>-abhängigen 3-Methylaspartat-Abbauwege werden von den Enzymen eines einzelnen Organismus durchgeführt. Die beiden übrigen Fermentationswege erfordern die Beteiligung von mindestens einem weiteren Organismus. Viele aerobe und anaerobe Organismen überführen das ubiquitäre Glutamat in andere Aminosäuren, die dann durch Clostridiales fermentiert werden. So ist beispielsweise *E. coli* in der Lage, Glutamat durch Decarboxylierung in den Neurotransmitter 4-Aminobutyrat umzuwandeln [Castanie-Cornet & Foster 2001; Castanie-Cornet *et al.* 1999]. Der weitere Abbau dieser  $\gamma$ -Aminosäure im entsprechenden Fermentationsweg konnte bisher nur für *Clostridium aminobutyricum* eindeutig nachgewiesen werden [Hardman & Stadtman 1960] und verläuft mit der folgenden Bilanz [Buckel 2001b]:



$$\Delta G^\circ = -37 \text{ kJ mol}^{-1} / \text{4-Aminobutyrat}; 74 \text{ kJ mol}^{-1} / \text{ATP}$$

Die Fermentation von Glutamat über 5-Aminovaleriat setzt sogar die Beteiligung von drei Organismen voraus: *Bacillus subtilis* reduziert unter osmotischem Stress Glutamat zu Prolin [Bremer & Krämer 2000], das von *Clostridium sporogenes* zu 5-Aminovaleriat weiterreduziert wird [Barker 1961; Stickland 1935]. Dieses wiederum wird schließlich von *Clostridium viride* entsprechend der folgenden Gleichung weiter abgebaut [Buckel *et al.* 1994]:



$$\Delta G^\circ = -37 \text{ kJ mol}^{-1} / \text{5-Aminovaleriat}; 74 \text{ kJ mol}^{-1} / \text{ATP}$$

Neben diesen fünf Fermentationswegen gibt es zwei weitere Abbauwege, die nur in Syntrophie mit methanogenen Archaeen auftreten. Diese sind in der Lage, den  $\text{H}_2$ -Partialdruck auf 1 Pa zu reduzieren, und ermöglichen somit die Entladung der Reduktionsäquivalente durch die Bildung von elementarem Wasserstoff. In Kokultur mit dem Methanbildner *Methanobacterium thermoautotrophicum* Z245 wurden so für *Caloramator coolhaasii* und *Thermanaerovibrio acidaminovorans* Glutamat-Fermentationswege nachgewiesen, die zur Bildung von Acetat bzw. Propionat führen [Plugge *et al.* 2001] und bis zu 1.0 ATP/Glutamat liefern [Buckel 2001b].

Der klassische Weg der Glutamat-Fermentation verläuft über die Bildung von 3-Methylaspartat. Er wurde zunächst im Bodenbakterium *Clostridium tetanomorphum* entdeckt [Barker 1961] und konnte später auch in den nahe verwandten Bakterien *C. tetani*, *C. cochlearium*, *C. malenominatum*, *C. lentoputrescens* und *C. limosum* nachgewiesen werden [Buckel 1980]. Die erste und gleichzeitig ungewöhnlichste Reaktion dieses Abbauweges stellt die Umwandlung von (S)-Glutamat in (2S,3S)-3-Methylaspartat dar, die mit der Neuordnung des Kohlenstoffgerüsts einhergeht: Sie wird von dem  $\text{B}_{12}$ -abhängigen Enzym Glutamat-Mutase nach einem einzigartigen radikalischen Mechanismus katalysiert [Buckel & Golding 1998]. Der weitere Verlauf des Abbauweges resultiert in der Bildung von Ammonium,  $\text{CO}_2$ , Acetat, Butyrat und  $\text{H}_2$  (max.  $\text{H}_2$ -Partialdruck = 20 kPa):



$$\Delta G^\circ = -63 \text{ kJ mol}^{-1} / \text{Glutamat}; 106 \text{ kJ mol}^{-1} / \text{ATP}$$

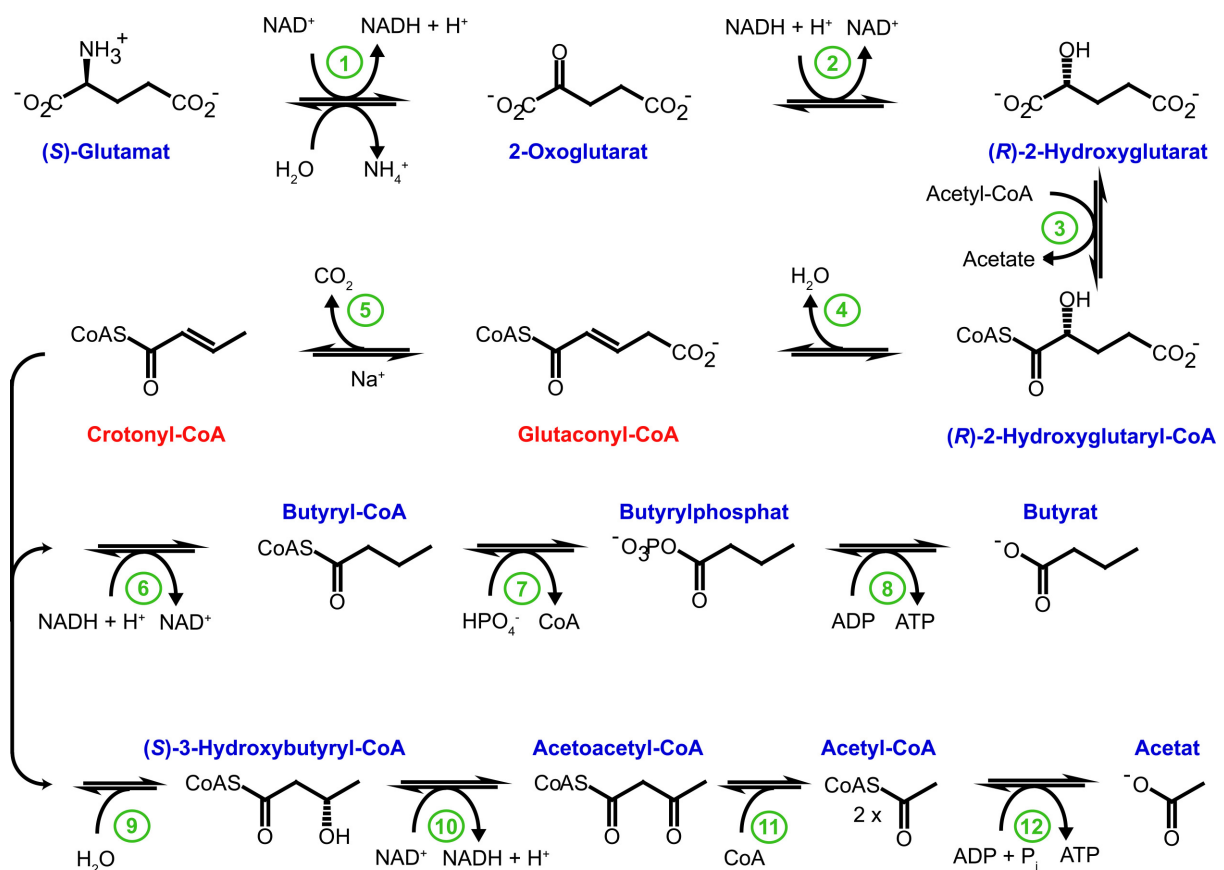
Dabei werden bis zu drei ATP-Moleküle durch Substratkettenphosphorylierung über Acyl-Phosphate gebildet. Dies entspricht einer Ausbeute von 0.6 ATP/Glutamat.

Eine Variante des 3-Methylaspartat-Weges, bei der Pyruvat zu Acetat,  $\text{CO}_2$  und Propionat disproportioniert, wurde in *Selenomonas acidaminophila* und *Barkera propionica* entdeckt [Buckel 2001b; Nanninga *et al.* 1987]:



$$\Delta G^\circ = -69 \text{ kJ mol}^{-1} / \text{Glutamat}; 62 \text{ kJ mol}^{-1} / \text{ATP}$$





**Abbildung 1-8 Glutamat-Fermentation über (R)-2-Hydroxyglutarat** Die grünen Zahlen kennzeichnen die jeweiligen Enzyme: **1**, Glutamat-Dehydrogenase (EC 1.4.1.2); **2**, 2-Hydroxyglutarat-Dehydrogenase (EC 1.1.99.2); **3**, Glutaconat-CoA-Transferase (EC 2.8.3.12); **4**, 2-Hydroxyglutaryl-CoA-Dehydratase + Aktivatorprotein (EC 4.2.1.-); **5**, Glutaconyl-CoA-Decarboxylase (EC 4.1.1.70); **6**, Butyryl-CoA-Dehydrogenase (EC 1.3.99.2) + ETF (engl., *electron transferring flavoprotein*); **7**, Phosphat-Butyryl-Transferase (EC 2.3.1.19); **8**, Butyrat-Kinase (EC 2.7.2.7); **9**, 3-Hydroxybutyryl-CoA-Dehydratase (Crotonase, EC 4.2.1.55); **10**, 3-Hydroxybutyryl-CoA-Dehydrogenase (EC 1.1.1.157); **11**, Acetyl-CoA-C-Acetyl-Transferase (Thiolase, EC 2.3.1.9); **12**, Phosphat-Acetyl-Transferase (EC 2.3.1.8) und Acetat-Kinase (EC 2.7.2.1).

## Der 2-Hydroxyglutarat-Weg

In Mikroorganismen wie *Peptostreptococcus asaccharolyticus*, *Acidaminococcus fermentans*, *C. sporosphaeroides*, *C. symbiosum* und *Fusobacterium nucleatum*, die in anoxischen Nischen in Menschen und Tieren leben, erfolgt die Glutamat-Fermentation über den gut charakterisierten 2-Hydroxyglutarat-Weg [Buckel 1980; Buckel & Barker 1974]. Dieser verläuft nach der gleichen Nettoreaktion wie der bereits erwähnte 3-Methylaspartat-Weg. Es werden also mit einer Energieausbeute von  $\Delta G^\circ = -63$  kJ Ammonium, CO<sub>2</sub>, Acetat, Butyrat und Wasserstoff gebildet, wobei 106 kJ für 1 Mol ATP aufgewendet werden müssen. Die einzelnen Reaktionsschritte des vollständigen Abbaupfades sind in Abbildung 1-8 gezeigt.

Die Fermentation beginnt mit der  $\text{NAD}^+$ -abhängigen Desaminierung von Glutamat zu 2-Oxoglutarat durch die Glutamat-Dehydrogenase. Anschließend katalysiert die 2-Hydroxyglutarat-Dehydrogenase unter Bildung von  $\text{NAD}^+$  die stereoselektive Reduktion zu (R)-2-Hydroxyglutarat, das dann zum Thiolester (R)-2-Hydroxyglutaryl-CoA umgewandelt wird. Die beiden darauf folgenden Schritte sind vom mechanistischen und bioenergetischen Standpunkt aus betrachtet die interessantesten Reaktionen dieses Stoffwechselweges. Die stereochemisch äußerst anspruchsvolle *syn*-Eliminierung von Wasser bei der Dehydrierung von (R)-2-Hydroxyglutaryl-CoA zu Glutaconyl-CoA erfolgt über die Bildung eines Ketylradikal-Anions und wird durch das aus mehreren Komponenten bestehende  $[\text{4Fe-4S}]^{2+}$ -Cluster-Protein 2-Hydroxyglutaryl-CoA-Dehydratase katalysiert [Buckel 1980; Hans *et al.* 2000; Locher *et al.* 2001; Schweiger *et al.* 1987]. Die anschließende Decarboxylierung zu Crotonyl-CoA erfolgt durch das Enzym Glutaconyl-CoA-Decarboxylase, welches Gegenstand der vorliegenden Arbeit ist und im Folgenden noch ausführlicher behandelt werden wird. Die weitere Umsetzung von Crotonyl-CoA verzweigt sich in einen oxidativen und einen reduktiven Weg und führt schließlich zur Bildung von Acetat bzw. Butyrat.

Die Energiebilanz des 2-Hydroxyglutarat-Fermentationsweges und der Butyrat-Variante des 3-Methylaspartat-Weges liegt mit  $106 \text{ kJ mol}^{-1}/\text{ATP}$  deutlich höher als bei der Mehrheit der bekannten anaeroben und aeroben Stoffwechselwege, bei denen meist nur ca. 70-80 kJ aufgewendet werden müssen, um 1 Mol ATP zu erzeugen [Thauer *et al.* 1977]. Es stellt sich daher die Frage, in welcher Form diese Energiedifferenz von ca. 30 kJ verbraucht bzw. gespeichert wird. Im Falle des 3-Methylaspartat-Weges wurde der membrangebundene Enzymkomplex  $\text{NADH:Ferredoxin-Oxidoreduktase}$  identifiziert, der die exergone Reduktion von  $\text{NAD}^+$  durch reduziertes Ferredoxin katalysiert und möglicherweise in der Lage ist, die gewonnene Energie in Form eines  $\text{Na}^+$ -Ionen-Gradienten zu speichern [Boiangiu *et al.* 2005]. In der Glutamat-Fermentation über 2-Hydroxyglutarat konnte die Bildung eines solchen Gradienten nachgewiesen werden. Sie erfolgt durch das bereits erwähnte Membran-assoziierte Enzym Glutaconyl-CoA-Decarboxylase [Buckel & Semmler 1982]: Bei der Decarboxylierung von Glutaconyl-CoA zu Crotonyl-CoA werden  $\text{Na}^+$ -Ionen aus dem Inneren der Zelle über die Zytoplasmamembran gepumpt, wobei ein extrazelluläres Proton verbraucht wird. Die auf diese Weise gespeicherte Energie steht für endergone Reaktionen des Organismus zur Verfügung.

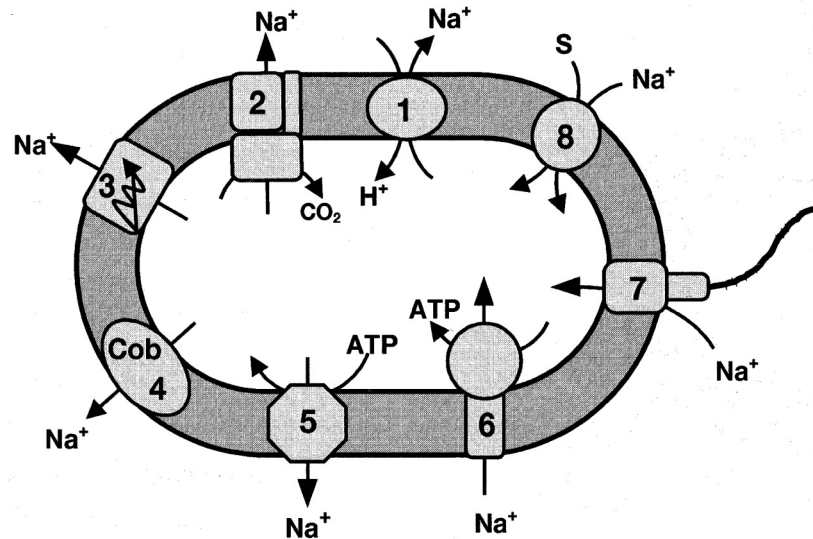
### 1.2.2 Die Rolle von Na<sup>+</sup> in der Kopplung bioenergetischer Prozesse

Der elektrochemische Protonengradient über die Membranen von Bakterien, Mitochondrien und Chloroplasten, oft als *PMF* abgekürzt (engl., *proton-motive force*), stellt die primäre Energiequelle für eine Vielzahl zellulärer Prozesse in den meisten Pro- und Eukaryoten dar. Der Protonen-Zyklus beinhaltet dabei typischerweise die Speicherung der PMF durch primäre Transportsysteme (H<sup>+</sup>-Pumpen) und ihren Verbrauch in der ATP-Synthese, in Transportvorgängen über die Zellmembran, in der Zellmigration usw. [Cramer & Knaff 1990; Skulachev 1988].

In Bakterien müssen die Protonen aus der Zelle herausgepumpt werden, so dass ihre Konzentration innerhalb der Zelle geringer ist als außerhalb. Je nach den Bedingungen, die ein Mikroorganismus in seiner Umgebung vorfindet, kann es schwierig sein, eine ausreichend hohe PMF aufzubauen. Dies gilt besonders für Bakterien, die in ökologischen Nischen leben, in denen die äußere Konzentration von H<sup>+</sup>-Ionen gering ist. So konnte nachgewiesen werden, dass Bakterien, die in alkalischem Milieu leben, statt Protonen zusätzlich oder sogar ausschließlich Natriumionen für die Kopplung bioenergetischer Prozesse verwenden [Dimroth 1990; Krulwich 1995; Skulachev 1985; Skulachev 1989]. Skulachev prägte dafür den Begriff der „Natriumwelt“ [Skulachev 1988]. In Analogie zur PMF spricht man hierbei von der *sodium-motive force*, kurz *SMF*.

Anfangs dachte man, dass sich die Bedeutung der sodium-motive force vor allem auf extremophile Organismen beschränken würde. Allerdings wurde das Konzept der SMF zunächst für anaerobe Bakterien wie *Propionigenium modestum*, *Malonomonas rubra* und *Caloramator fervidus* nachgewiesen, die in ihrem Energiemetabolismus ausschließlich Natriumionen nutzen und keineswegs Extremophile sind [Dimroth & Hilbi 1997; Hilpert *et al.* 1984; Speelmans *et al.* 1993]. Darüber hinaus zeigen neuere Untersuchungen, dass die Genome einer Reihe von alkaliphilen und hyperthermophilen Organismen keine der bekannten primären Na<sup>+</sup>-Pumpen enthalten [Deckert *et al.* 1998; Falb *et al.* 2005; Kawarabayasi *et al.* 2001; She *et al.* 2001; Takami *et al.* 2000]. Dagegen konnte für einige mesophile Bakterien, darunter verschiedene marine [Kogure 1998] und mehrere humanpathogene Stämme [Hase *et al.* 2001; Zhou *et al.* 1999] gezeigt werden, dass ihr Energiehaushalt Na<sup>+</sup>-gekoppelt ist. Es ist nicht völlig klar, warum einige Mikroorganismen SMF, PMF oder beide anwenden, allerdings wird vermutet, dass die ursprüngliche Form der Energiespeicherung durch Ionengradienten zunächst aus-

schließlich auf  $\text{Na}^+$ -Ionen basierte und somit die SMF der Vorläufer der PMF ist [Castanie-Cornet *et al.* 1999; Mulkidjanian *et al.* 2008a; Mulkidjanian *et al.* 2008b].



**Abbildung 1-9** Schematische Darstellung  $\text{Na}^+$ -Ionen-gekoppelter bioenergetischer Prozesse [Dimroth *et al.* 2001] Erläuterungen siehe Text

In Abbildung 1-9 ist eine Auswahl bioenergetischer Prozesse in Bakterien dargestellt, die über die Zirkulation von Natriumionen gekoppelt sind [Dimroth *et al.* 2001]. Die SMF stellt dabei die Energie zur Verfügung, die benötigt wird, um zelluläre Prozesse wie z.B. (6) die Gewinnung von ATP durch eine  $F_1F_0$ -ATP-Synthase, (7) den Antrieb eines Flagellenmotors oder (8) Transportprozesse zur Aufnahme von Nährstoffen anzutreiben. Für die Erzeugung der SMF kommen eine Reihe von verschiedenen Enzymen und Enzymkomplexen in Frage. Dazu gehören Enzyme der (3) Atmungskette, wie beispielsweise die NADH:Ubiquinone-Oxidoreduktase aus *Vibrio alginolyticus* [Hayashi *et al.* 2001; Pfenninger-Li *et al.* 1996], *Klebsiella pneumoniae* [Krebs *et al.* 1999] oder *Escherichia coli* [Steuber *et al.* 2000], oder (4)  $\text{Na}^+$ -pumpende Methyltetrahydromethanopterin:CoM-Methyltransferasen aus methanogenen Mikroorganismen [Becher *et al.* 1992; Gartner *et al.* 1994]. Darüber hinaus gibt es (1)  $\text{Na}^+/\text{H}^+$ -Antiporter, die die  $\text{Na}^+$ - und  $\text{H}^+$ -Zirkulation miteinander koppeln [Padan & Schuldiner 1993], und (5) V-Typ-ATPasen, die  $\text{Na}^+$ -Ionen unter ATP-Hydrolyse aus der Zelle pumpen. Schließlich wurden in einigen Organismen Biotin-abhängige Decarboxylasen identifiziert, denen es möglich ist, die frei werdende Energie der von ihnen katalysierten Reaktionen

in Form von  $\text{Na}^+$ -Gradienten zu speichern [Dimroth 1997]. Zu dieser Klasse Natriumionen-pumpender Decarboxylasen gehört auch die bereits erwähnte Glutaconyl-CoA-Decarboxylase.

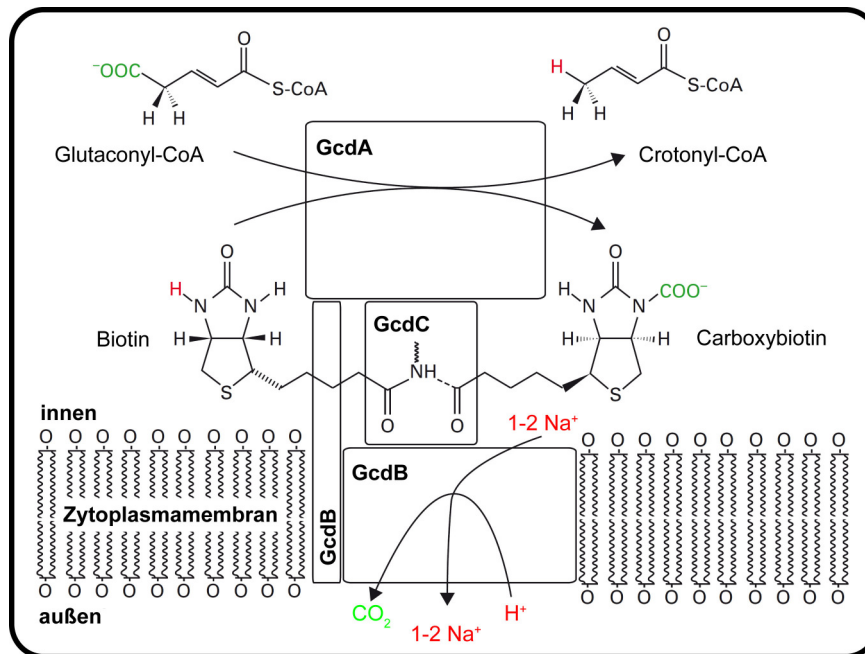
### 1.2.3 Die Natriumionen-Pumpe Glutaconyl-CoA-Decarboxylase

$\text{Na}^+$ -pumpende Decarboxylasen sind Biotin-abhängige Enzyme, die aus mehreren zytosolischen und membranständigen Untereinheiten aufgebaut sind und die Decarboxylierung von  $\beta$ -Ketosäuren oder analogen Verbindungen katalysieren. Sie sind in der Lage, die frei werdende Energie dieser Reaktion zu nutzen, um  $\text{Na}^+$ -Ionen vom Zytoplasma aus der Zelle herauszutransportieren und somit einen  $\text{Na}^+$ -Ionen-Gradienten aufzubauen, der Energie für endergone zelluläre Prozesse liefert. Die Oxalacetat-Decarboxylase aus *Klebsiella aerogenes* war das erste Mitglied dieser Enzymfamilie, für das nachgewiesen werden konnte, dass es als primäre  $\text{Na}^+$ -Pumpe fungiert [Dimroth 1982]. Mittlerweile wurde eine ganze Reihe ähnlicher Enzyme in weiteren Mikroorganismen identifiziert [Buckel 2001a]. Oxalacetat-Dehydrogenase (EC 4.1.1.3, Oad), Methylmalonyl-Decarboxylase (EC 4.1.1.41, Mmd) und Glutaconyl-CoA-Decarboxylase (EC 4.1.1.70, Gcd) sind die am besten charakterisierten Vertreter dieser Klasse von Enzymen.

#### Zusammensetzung des Enzymkomplexes

Die erste Natriumionen-pumpende Glutaconyl-CoA-Decarboxylase wurde bereits 1982 von Buckel und Semmler in dem strikt anaeroben Darmbakterium *Acidaminococcus fermentans* nachgewiesen [Buckel & Semmler 1982]. Das Enzym kann auf Grund seines Biotin-Gehalts problemlos direkt aus solubilisierten Membranfraktionen mittels Avidin-Affinitätschromatographie aufgereinigt werden. Das in Phospholipid-Vesikeln rekonstituierte Enzym zeigt eine eindeutig nachweisbare  $\text{Na}^+$ -Pumpen-Aktivität: Die Zugabe des Kaliumionophors Valinomycin beschleunigt in Gegenwart von Kaliumionen den  $\text{Na}^+$ -Transport, während er in Gegenwart des  $\text{Na}^+$ -Ionophors Monensin vollständig unterbunden wird [Buckel & Semmler 1982]. Durch SDS-PAGE-Analyse und Edman-Sequenzierung des aufgereinigten Enzyms wurden vier unterschiedliche Typen von Untereinheiten,  $\alpha$ ,  $\beta$ ,  $\gamma$  und  $\delta$ , identifiziert, die im Folgenden auch als GcdA, GcdB, GcdC bzw. GcdD bezeichnet werden und berechnete Molekulargewichte von ca. 65, 35, 14 bzw. 12 kDa aufweisen [Braune *et al.* 1999]. Der Aufbau des Glutaconyl-CoA-Decarboxylase-Enzymkomplexes und das Zusammenspiel der Un-

tereinheiten bei der Katalyse sind in Abbildung 1-10 schematisch dargestellt [Boiangiu *et al.* 2005].



**Abbildung 1-10 Schematische Darstellung des Aufbaus des Glutaconyl-CoA-Enzymkomplexes zur Erläuterung der Funktion und des Zusammenspiels der Untereinheiten [Boiangiu *et al.* 2005]**

Die hydrophile Decarboxylase-Untereinheit GcdA katalysiert den  $\text{Na}^+$ -unabhängigen Carboxylat-Transfer vom CoA-Ester-Substrat Glutaconyl-CoA auf Biotin. Letzteres ist kovalent an das Biotin-Carrier-Protein GcdC gebunden. Bei der hydrophoben Untereinheit GcdB handelt es sich um ein integrales Membranprotein mit 9-11 Transmembranhelices [Boiangiu *et al.* 2005; Buckel 2001a]. Dieses katalysiert die Decarboxylierung des gebildeten  $\text{N}^1$ -Carboxybiotins und nutzt die dabei frei werdende Energie, um unter Verbrauch eines extrazellulären Protons 1-2  $\text{Na}^+$ -Ionen aus dem Inneren der Zelle über die Zytoplasmamembran nach außen zu transportieren [Buckel 2001a]. Die Funktion der kleinen  $\delta$ -Untereinheit konnte noch nicht ermittelt werden. Sie enthält wahrscheinlich eine Transmembranhelix und trägt möglicherweise als Membrananker für GcdA zur Stabilität des Komplexes bei [Braune *et al.* 1999]. Analog zur *A. fermentans*-Gcd wurden die Glutaconyl-CoA-Decarboxylasen aus den anaeroben Darmbakterien *Peptostreptococcus asaccharolyticus* und *Clostridium symbiosum* [Buckel & Semmler 1983] sowie dem gram-negativen Plague-Bakterium *Fusobacterium nucleatum* subsp. *nucleatum* [Beatrix *et al.* 1990] erfolgreich als primäre Natriumionen-Pumpen charakterisiert.

Abbildung 1-11A zeigt eine SDS-PAGE-Analyse der aufgereinigten Glutaconyl-CoA-Decarboxylasen aus *A. fermentans*, *C. symbiosum* und *F. nucleatum*, die die generelle Zusammensetzung des Enzymkomplexes bestätigt. Der einzige signifikante Unterschied zu *A. fermentans* ist das Verhalten der Biotin-Carrier-Untereinheiten: Während GcdC aus *A. fermentans* auf dem SDS-PAGE-Gel ein Muster von mehreren Banden bei einem apparenten Molekulargewicht von ca. 120 kDa zeigt, weist *F. nucleatum*-GcdC ein normales Migrationsverhalten mit einer einzelnen Bande bei ca. 19 kDa auf, wohingegen bei *C. symbiosum*-GcdC zwei Banden bei unterschiedlichem Molekulargewicht erkennbar sind (ca. 22 und 24 kDa). Diese Beobachtungen wurden jeweils durch Western-Blot-Analysen mit Hilfe eines Anti-Biotin-Antikörpers verifiziert. Während es sich beim ungewöhnlichen Verhalten der *A. fermentans*-GcdC wahrscheinlich um ein Artefakt der Gelelektrophorese handelt [Beatrix *et al.* 1990], belegen die Ergebnisse der Edman-Sequenzierung, dass *C. symbiosum*-Gcd tatsächlich zwei unterschiedliche Biotin-Carrier-Untereinheiten, GcdC<sub>1</sub> und GcdC<sub>2</sub>, enthält (Abbildung 1-12) [Kress *et al.* 2009].

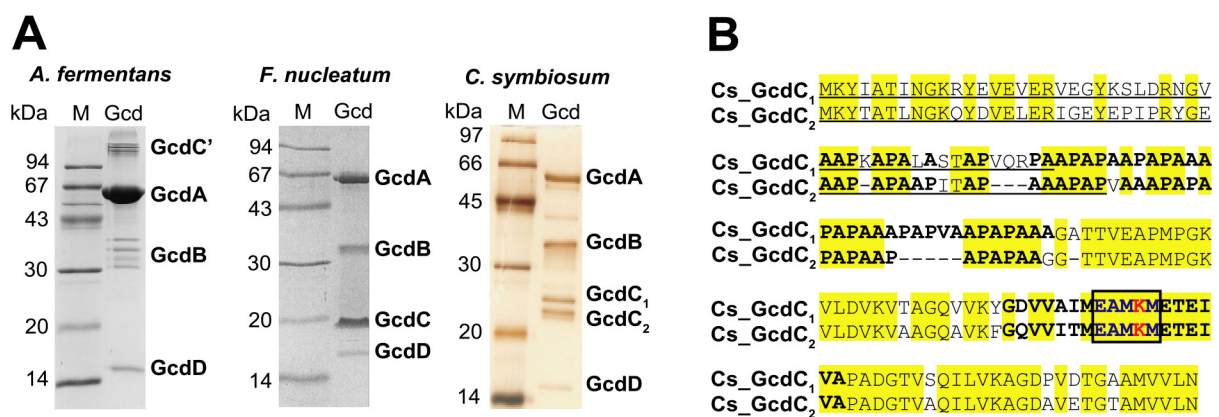
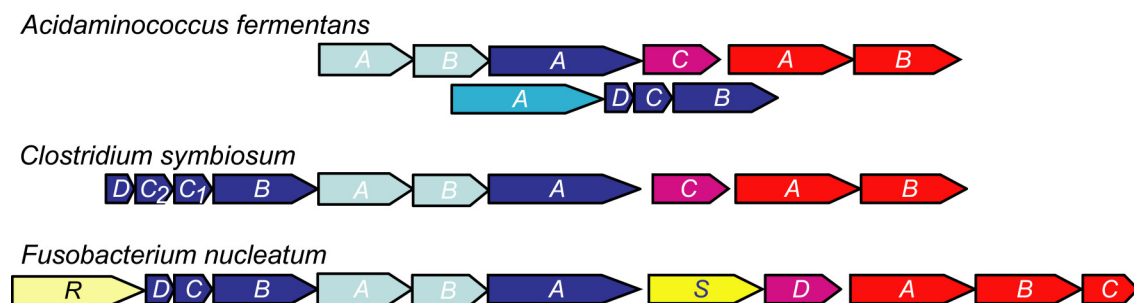


Abbildung 1-11 (A) SDS-PAGE-Analyse der aufgereinigten Glutaconyl-CoA-Decarboxylasen aus *A. fermentans* [Braune *et al.* 1999], *F. nucleatum* [Beatrix *et al.* 1990] und *C. symbiosum* [Kress *et al.* 2009] (M  $\triangleq$  Massenstandard) (B) Vergleich der Sequenzen von GcdC<sub>1</sub> und GcdC<sub>2</sub> aus *C. symbiosum* [Kress *et al.* 2009; Zelder 2008] Die Sequenzierung erfolgte über den Edman-Abbau der jeweiligen N-terminalen Domäne (unterstrichen). Konservierte Reste sind gelb markiert. Fett gedruckte Aminosäuren heben die AP-reiche Linker-Region und das Biotin-Bindemotiv (EAMKM) hervor. Dieses in Biotin-abhängigen Decarboxylasen hochkonservierte Motiv ist zusätzlich schwarz umrandet und der Lysin-Rest, an den das Biotin gebunden ist, in rot dargestellt.

Die beiden GcdC-Untereinheiten haben eine Sequenzidentität von 79% (Abbildung 1-11B) und weisen berechnete molekulare Massen von 15 bzw. 14 kDa auf. Ihr allgemeiner Aufbau entspricht dem der Biotin-Carrier-Proteine aus *A. fermentans* und *F. nucleatum*. Die

C-terminalen Regionen von GcdC<sub>1</sub> und GcdC<sub>2</sub> (A78-N149 bzw. A68-N139) beinhalten das konservierte Biotin-Bindemotiv (EAMKM) und stellen somit die eigentliche Biotin-Carrier-Domäne dar, während die N-terminale Domäne (M1-A31) wahrscheinlich der Verankerung im Enzymkomplex dient [Wendt *et al.* 2003]. Der Hauptunterschied zwischen den beiden GcdC-Untereinheiten ist die unterschiedliche Länge des flexiblen Linker-Bereichs, der die beiden Domänen verbindet und größtenteils aus Alanin- und Prolin-Resten besteht.



**Abbildung 1-12 Organisation der für Glutaconyl-CoA-Decarboxylasen codierenden Gene [Kress *et al.* 2009]** Dargestellt sind die Gencluster der Glutaconyl-CoA-Decarboxylase-Gene (*gcd*, dunkelblau) in *A. fermentans*, *C. symbiosum* und *F. nucleatum*. Die weiteren Gene codieren für Glutaconat-CoA-Transferase (*gct*, hellblau) und 2-Hydroxyglutaryl-CoA-Dehydratase (*hgd*, rot) sowie für deren Aktivator (*hgdC* bzw. *hgdD*, magenta) [Hans *et al.* 1999]. Die zusätzlichen Gene in *F. nucleatum* codieren für den Regulator *R* und den Glutamat/Na<sup>+</sup>-Symporter *S* (beide in gelb). In *A. fermentans* bilden die Gene der kleineren Glutaconyl-CoA-Decarboxylase-Untereinheiten D, C und B zusammen mit dem Gen der Untereinheit A der Methylmalonyl-CoA-Decarboxylase (*mmdA*, türkis) ein separates Operon.

## Mechanismus der Decarboxylase-Reaktion

Auf der Grundlage ihrer Kristallstruktur ordneten Benning *et al.* im Jahr 2000 das Biotin-unabhängige Enzym *E. coli*-Methylmalonyl-CoA-Decarboxylase der Crotonase-Strukturfamilie zu [Benning *et al.* 2000; Haller *et al.* 2000] und formulierten einen zu 4-Chlorobenzoyl-CoA-Dehalogenase (EC 3.8.1.6, [Benning *et al.* 1996]) und Enoyl-CoA-Dehydratase (Crotonase, EC 4.1.2.17, [Engel *et al.* 1996]) analogen Reaktionsmechanismus. Unter der Annahme, dass die Bindung von Glutaconyl-CoA an die Decarboxylase-Untereinheit der Gcd in ähnlicher Weise erfolgt wie bei diesen Vertretern der Crotonase-Familie, wurde für Glutaconyl-CoA-Decarboxylase der in Abbildung 1-13 illustrierte Mechanismus vorgeschlagen [Buckel 2001a].



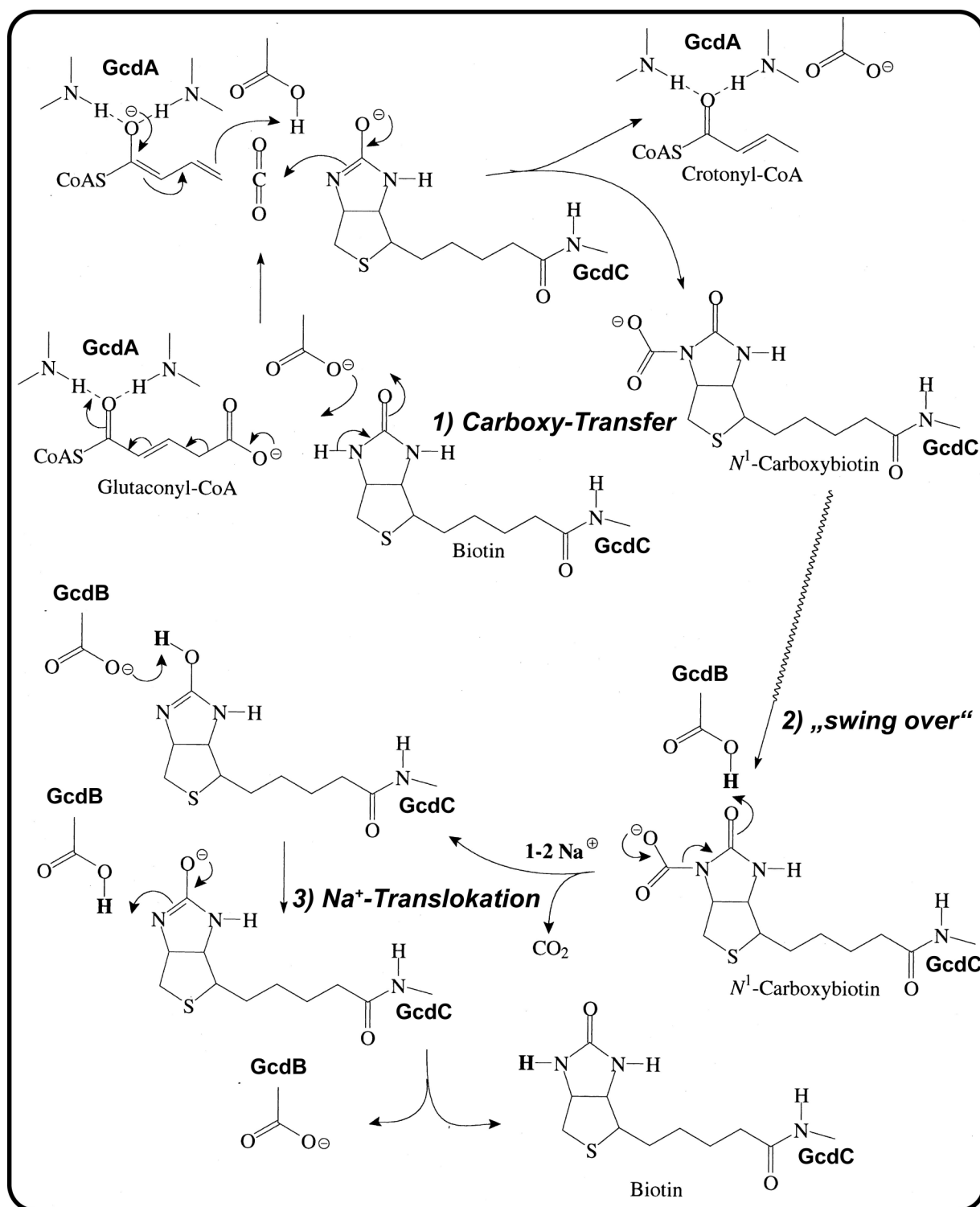


Abbildung 1-13 Hypothetischer Mechanismus der Glutaconyl-CoA-Decarboxylase in Anlehnung an den Mechanismus der Enoyl-CoA-Hydratase (Crotonase) [Buckel 2001a] Die Gesamtreaktion läuft in drei Phasen ab: 1. Übertragung der Carboxy-Gruppe von Glutaconyl-CoA auf Biotin (Carboxy-Transfer), 2. Transfer des N<sup>1</sup>-Carboxybiotins zur GcdB-Untereinheit („swing over“) 3. Decarboxylierung des N<sup>1</sup>-Carboxybiotins und gleichzeitige Na<sup>+</sup>-Translokation.

Demnach wird das im ersten Schritt der Reaktion (*Carboxy-Transfer*) durch spontane Decarboxylierung von Glutaconyl-CoA gebildete Dienolat durch Wasserstoffbrückenbindung mit zwei NH-Gruppen des Protein-Rückgrats stabilisiert. Gleichzeitig entfernt eine basische Seitenkette das N<sup>1</sup>-Proton von Biotin und überträgt es auf das Dienolat, wobei die Methylgruppe von Crotonyl-CoA entsteht. Mit der Bildung von N<sup>1</sup>-Carboxybiotin durch nukleophilen Angriff des Biotin-Anions auf CO<sub>2</sub> wird dieser Teil der Reaktion abgeschlossen. Im nächsten Schritt („*swing over*“) wird das Carboxybiotin wahrscheinlich durch eine konformationelle Änderung der GcdC-Untereinheit zur Na<sup>+</sup>-Pumpe GcdB transferiert. Die sich anschließende Decarboxylierung erfolgt durch Protonierung des 2-Carbonyl-Sauerstoffs des Carboxybiotins, wobei gleichzeitig 1-2 Na<sup>+</sup>-Ionen auf die Außenseite der Zytoplasmamembran gepumpt werden (*Na<sup>+</sup>-Translokation*). Das dabei verbrauchte Proton stammt aus dem Periplasma bzw. von außerhalb der Zelle [Di Berardino & Dimroth 1996]. Für die Funktionsweise des Pumpenmechanismus der verwandten Natriumionen-Pumpe Oxalacetatdecarboxylase wurde von Dimroth et al. mit Hilfe von ortsgerichteter Mutagenese ein Modell abgeleitet, auf das an dieser Stelle nicht näher eingegangen werden kann [Dimroth et al. 2001].

Die Zugehörigkeit der Decarboxylase-Untereinheit der Gcd aus *A. fermentans* zur Crotonase-Familie konnte durch die Aufklärung ihrer Kristallstruktur bestätigt werden [Wendt et al. 2003]. Die Bindung des Glutaconyl-CoA-Esters erfolgt demnach tatsächlich in der beschriebenen Weise, allerdings wurden im aktiven Zentrum keine Aminosäurereste entdeckt, die die Funktion der postulierten Säure oder Base übernehmen könnten. Es wird daher angenommen, dass Biotin sein N<sup>1</sup>-Proton direkt auf das durch spontane Decarboxylierung entstehende Dienolat von Crotonyl-CoA überträgt [Boiangiu et al. 2005].

## 1.3 Zielsetzung

### **Enamidase**

Das Metall-abhängige Enzym Enamidase ist von Alhapel et al. als Schlüsselenzym der Nicotinat-Fermentation in *Eubacterium barkeri* identifiziert worden [Alhapel 2005; Alhapel et al. 2006]. Ziel der vorliegenden Arbeit war es, durch die röntgenographische Strukturanalyse des in *E. coli* rekombinant überproduzierten Enzyms einen Beitrag zur Aufklärung des Mechanismus der von ihm katalysierten außergewöhnlichen stereoselektiven Zweistufenreaktion zu leisten. Darüber hinaus sollte die Enamidase aus dem Organismus *Bradyrhizobium japonicum* kristallisiert und strukturell charakterisiert werden.

### **Glutaconyl-CoA-Decarboxylase**

Nachdem im Vorfeld Versuche, die aus *Clostridium symbiosum* aufgereinigte Glutaconyl-CoA-Decarboxylase als intakten Enzymkomplex zu kristallisieren, nicht zum Erfolg geführt hatten [Kress et al. 2009], sollten durch die Kristallisation und röntgenographische Strukturaufklärung der in *E. coli* rekombinant überproduzierten Decarboxylase-Untereinheit Einblicke in Substratbindung, Mechanismus und Quartärstruktur dieser primären Natriumionen-Pumpe gewonnen werden.



## 2 Material

### 2.1 Geräte, Verbrauchsmaterialien und Chemikalien

Tabelle 2-1 Geräte

Gerät	Typenbezeichnung (Hersteller)
Autoklaven	T-line™ (Fedegari), V-65 (Systec)
Brutschränke	Brutschrank (Heraeus), BFED-53 (WTB Binder)
Digitalkamera	Powershot A70 (Canon)
Dokumentationssystem	
Agarosegele	Computer E.A.S.Y. mit Kamera (UVP), UV-Schirm (Herolab), Thermodrucker UP-D 895 (Sony)
Kristallisationsansätze	Crystal Pro HT 110 (TriTek Corporation)
Fluidizer	EmulsiFlex-C5 (Avestin)
French Press	French-Pressure Cell-Version 5.1 (Sim Aminco)
Gelelektrophorese	
SDS-PAGE	Mighty Small II SE250 (HoeferScientific Instruments), PerfectBlue Doppelgelsystem Twin M (PEQLAB)
Spannungsquellen	Model 453 (ISCO), EPS 300 (GE Healthcare)
Gelschüttler	Modell 3013 (GFL), Köttermann 4010 (Köttermann)
Heizblöcke / Thermomixer	BT3 / UBD (Grant), Thermomixer compact (Eppendorf)
Kristallisationsroboter	Cartesian Microsys™ SQ4000 (Genomic Solutions)
LC-Systeme	AEKTAprime™ P100 (GE Healthcare), ÄKTApurifier™ 100 (GE Healthcare), FPLC System 250 (GE Healthcare)
Liquid handling-System	Lissy (Zinsser Analytic)
Mikroskope	B061 (Olympus), MZ8 (Leica)
pH-Meter	Basic Meter PB-11 (Sartorius), Seven Easy (Mettler Toledo)
Photometer	U-2000 (Hitachi), Ultrospec™ 3100pro (Biochrom), Uvikon Spectrophotometer 930 (Kontron), Nicolet Evolution e100 (Thermo Electron Corporation)

Reinstwasser-Anlage	Seralpur Pro90CN ( <i>Seral</i> ), Milli-Q ( <i>Millipore</i> )
Röntgendiffraktometer	Drehanode AXS FR591 ( <i>Bruker</i> ), MAR345dtb Detektor ( <i>MAR Research</i> ), Oxford Cryosystems ( <i>Oxford</i> )
Schlauchpumpe	Peristaltic pump P1 ( <i>GE Healthcare</i> )
Schüttelinkubatoren	Certomat® IS ( <i>Sartorius</i> ), G25 ( <i>New Brunswick Scientific</i> ), Multitron 2 ( <i>Infors</i> ), TH 25 ( <i>Edmund Bühler</i> )
Spülmaschine	professional G7883 ( <i>Miele</i> )
Ultraschallgerät	Sonoplus HD200 mit Spitze Typ UW 200 ( <i>Bandelin electronic</i> )
Waagen	DeltaRange® PC2200 ( <i>Mettler</i> ), Toledo Labstyle 54 ( <i>Mettler</i> )
Zentrifugen	2K15 ( <i>Sigma</i> ), Airfuge™ ( <i>Beckmann</i> ), Biofuge 13R ( <i>Heraeus</i> ), Biofuge fresco / pico ( <i>Heraeus</i> ), Centrifuge 5417C / 5910R ( <i>Eppendorf</i> ), Discovery™ 90SE ( <i>Sorvall</i> ), EvolutionRC ( <i>Sorvall</i> ) J2-HS / J2-21 M/E ( <i>Beckmann</i> ), Minifuge GL ( <i>Heraeus Christ</i> )

Tabelle 2-2 Verbrauchsmaterialien

Produkte	Typenbezeichnung ( <i>Hersteller</i> )
Dialysekassetten	Slide-A-Lyzer® Mini Dialysis Units 10000 MWCO ( <i>Pierce</i> )
Dialysemembran	Porengröße 0.025 µm, Ø = 25 mm ( <i>Millipore</i> ), SnakeSkin, MWCO 3.5 kDa ( <i>ThermoScientific</i> )
Einmalküvetten	Cuvettes REF 67.742 ( <i>Sarstedt</i> )
Entsalzungssäulen	PD-10 Desalting Columns ( <i>GE Healthcare</i> )
Kristallisationszubehör	
Platten	Deckgläser Ø = 22 mm ( <i>Hampton Research</i> ), 24 well Crystallization Tool ( <i>NeXtal/Qiagen</i> ), 24 well VDX ( <i>Hampton Research</i> ), 96 well CrystalQuick™ Platte ( <i>Greiner Bio-one</i> ), 96 well Innovaplate™ SD-2 ( <i>Hampton Research</i> ), SilverSeal ( <i>Greiner Bio-one</i> ), VIEWseal™ Klebefolien ( <i>Greiner Bio-one</i> )
Screens	NEXTAL Crystallization Kits ( <i>Qiagen</i> ), Protein Crystallization Kits ( <i>Sigma-Aldrich</i> ), Additiv-Screens ( <i>Sigma-Aldrich</i> , <i>Hampton Research</i> )
Werkzeuge	Cryozangen, CrystalWand™, Micro-Tools™ (alles <i>Hampton Research</i> )

Cryoschleifen	CrystalCap™ Coloured ( <i>Hampton Research</i> ), CrystalCap HT™ ( <i>Hampton Research</i> ), Mounted CryoLoop™ ( <i>Hampton Research</i> ), MicroMounts™ ( <i>MiTeGen</i> )
Proteinkonzentratoren	Amicon-Ultra 4 ( <i>Millipore</i> )
Sterilfilter	Ultrafree®-MC ( <i>Millipore</i> ), Filtropur S 0.2 ( <i>Sarstedt</i> ), Steritop Bottletop Filter, 22 µm Express plus ( <i>Millipore</i> )
Zubehör für LC-Systeme	Gelfiltrations-Kalibrierkit ( <i>GE Healthcare</i> ), Säulenkörper und -materialien ( <i>GE Healthcare</i> )

Tabelle 2-3 Chemikalien und Enzyme

Produkte	Hersteller
Acrylamid (Rothiphorese® Gel 30)	<i>Roth</i>
Agar	<i>Roth</i>
Ammoniumperoxodisulfat (APS)	<i>Fluka</i>
Ampicillin	<i>Applichem</i>
Bacto®-Hefeextrakt	<i>Difco</i>
Bacto®-Trypton	<i>Difco</i>
BCA-Assay	<i>Uptima</i>
Biocytin	<i>Sigma-Aldrich</i>
Carbenicillin	<i>Applichem</i>
Chloramphenicol	<i>Applichem</i>
Coomassie-Brilliant Blue G250	<i>Serva</i>
D-Desthiobiotin	<i>IBA</i>
n-Dodecyl-β-D-maltosid (DDM)	<i>GLYCON Biochemicals</i>
n-Decyl-β-D-maltosid (DM)	<i>GLYCON Biochemicals</i>
Dithiothreitol (DTT)	<i>Biomo</i>
Ethylendiamintetraessigsäure (EDTA)	<i>Roth</i>
Ethanol	<i>Riedel de Haën</i>
Glycerin	<i>KMF optichem</i>

2-(4-Hydroxyphenylazo)-benzoesäure (HABA)	<i>Fluka</i>
Isopropyl- $\beta$ -D-thiogalacto-pyranosid (IPTG)	<i>Duchefa</i>
Lysozym	<i>Fluka</i>
Natriumdodecylsulfat (SDS)	<i>Merck</i>
$\beta$ -Mercaptoethanol	<i>Roth</i>
Phenylmethylsulfonylfluorid (PMSF)	<i>Fluka</i>
peqGOLD Protein-Marker I	<i>PEQLAB</i>
Tetramethylenethyldiamin (TEMED)	<i>Sigma-Aldrich</i>
Virkon®	<i>Antec International</i>
<i>Strep</i> -Tactin MacroPrep Cartridge H-PR	<i>IBA</i>

---

Alle in der Tabelle nicht gesondert aufgeführten Chemikalien wurden in p.a.-Qualität von den Firmen *Merck*, *Sigma*, *Roth* bzw. *Fluka* bezogen.

Für Kokristallisations- und *soaking*-Experimente (vgl. 3.3.3) mit Enamidase und GcdA wurden die Verbindungen 1,4,5,6-Tetrahydro-6-oxonicotinat (THON), 6-Oxo-1,4,5,6-tetrahydropyridazin, 6-Hydroxynicotinat und 2-(Hydroxymethyl)-glutarat bzw. Glutaconyl-CoA, Crotonyl-CoA und Glutaryl-CoA verwendet. Diese wurden von den Arbeitsgruppen von Wolfgang Buckel (Mikrobiologie, Universität Marburg; GcdA) bzw. Antonio Pierik (Zytobiologie, Universität Marburg; Enamidase) zur Verfügung gestellt.



## 2.2 Vektoren und Mikroorganismen

Für die heterologe Expression des *gcdA*-Gens wurde der Vektor pASK-IBA7plus (IBA) in Verbindung mit dem *E. coli*-Stamm Rosetta™ (DE3)pLysS (Novagen) verwendet.

### pASK-IBA7plus

Dieser Vektor wurde für die cytosolische Expression rekombinanter Proteine in entsprechenden *E. coli*-Stämmen konzipiert und trägt eine Ampicillin-Resistenz-Kassette. Die Expression des Zielgens steht unter Kontrolle des Tetracyclin-Promotors (tet-Promotor). Das tet-Repressor-Gen befindet sich ebenfalls auf dem Plasmid und wird konstitutiv vom  $\beta$ -Lactamase-Promotor exprimiert. Auf diese Weise ist eine stringente Repression gewährleistet. Der tet-Promotor ist im Gegensatz zum lac-Promotor unempfindlich gegenüber dem Einfluss von Katabolit-Repression (z.B. cAMP-Konzentration) und chromosomal exprimierten Repressor-Molekülen. Die Induktion der Expression erfolgt durch Zugabe geringer Mengen Anhydrotetracycline.

Des Weiteren bietet der Vektor die Möglichkeit, das Zielgen als N-terminales Strep-tag II-Fusionsprotein zu exprimieren und damit die Aufreinigung durch Affinitätschromatographie zu vereinfachen.

### Rosetta™ (DE3)pLysS [F<sup>-</sup> *ompT hsdS<sub>B</sub>(r<sub>B</sub> m<sub>B</sub>) gal dcm* (DE3) pLysSRARE (Cam<sup>R</sup>)]

Dieser *E. coli*-Stamm enthält ein Chloramphenicol-Resistenz vermittelndes Plasmid, das für die tRNA der seltenen Codons AGG, AGA, AUA, CUA, CCC, GGA codiert. Dieses Plasmid trägt außerdem das T7-Lysozym-Gen, welches die basale Expression des T7-RNA-Polymerase-Gens zusätzlich reprimiert. Das Kürzel DE3 zeigt an, dass der Stamm das Lyso-Gen IDE3 und damit eine chromosomale Kopie der T7-RNA-Polymerase unter der Kontrolle des lacUV5-Promotors trägt.



## 2.3 Nährmedien, Stammlösungen und Puffer

Als Standard-Kulturmedium für die heterologe Genexpression diente LB-Medium. Für die Herstellung von Kulturplatten wurden vor dem Autoklavieren (30 min bei 121 °C, 1.5 bar) 15 g Agar in einen Liter flüssiges Kulturmedium gegeben. Für Selektionsexperimente auf antibiotikahaltigem Nährmedium wurden nach dem Autoklavieren Carbenicillin bzw. Chloramphenicol in den Endkonzentrationen 50 µg/mL bzw. 35 µg/mL zugesetzt.

LB-Medium	
10 g/L	Bacto®-Trypton
5 g/L	Bacto® Yeast Extract
10 g/L	NaCl
0.4 mL/L	10 N NaOH-Lösung

In den nachfolgenden Tabellen finden sich die Zusammensetzungen der in dieser Arbeit häufig verwendeten Stammlösungen und Puffer.

### Diverse

Antibiotika-Stammlösungen	
50 mg/mL	Carbenicillin (in ddH <sub>2</sub> O)
35 mg/mL	Chloramphenicol (in Isopropanol)

PMSF-Stammlösung (in Isopropanol)	
0.2 M Phenylmethylsulfonylfluorid	

EDTA-Stammlösung (in ddH <sub>2</sub> O)	
0.1M Ethylendiamintetraessigsäure	

Lysozym-Stammlösung (ddH <sub>2</sub> O)	
50mg/mL Lysozym	

**Stammlösungen für die SDS-PAGE**

Acrylamidlösung für SDS-Gele	
30 % (w/v)	Acrylamid
8 % (w/v)	Bisacrylamid
Sammelgelpuffer	
0.625 M	Tris-HCl, pH 6.8
Trenngelpuffer	
1.125 M	Tris-HCl, pH 8.0
30 % (w/v)	Saccharose
Laufpuffer	
0.25 M	Tris-Base
2 M	Glycin
1 % (w/v)	SDS
Probenpuffer	
62.5 mM	Tris-HCl, pH 6.8
15 % (w/v)	Glycerin
4 % (v/v)	SDS
0.001 % (w/v)	Bromphenolblau
4 % (v/v)	$\beta$ -Mercaptoethanol
62.5 mM	Tris-HCl, pH 6.8

Färbelösung	
0.25 % (w/v)	Coomassie Brilliant Blue R250
5 % (v/v)	Essigsäure
50 % (v/v)	Ethanol
1 h rühren, anschließend filtrieren	

### Puffer für die Affinitätschromatographie

Puffer W	
100 mM	Tris-HCl pH 8.0
150 mM	NaCl

Puffer E	
100 mM	Tris-HCl pH 8.0
150 mM	NaCl
2.5 mM	Desthiobiotin

Puffer R	
100 mM	Tris-HCl pH 8.0
150 mM	NaCl
1 mM	HABA



## 3 Methoden

### 3.1 Molekularbiologische Methoden

Die für die vorliegende Arbeit relevanten molekularbiologischen Methoden umfassten die Sequenzierung der die Zielproteine codierenden DNA-Fragmente, die Konstruktion und Sequenzierung der Expressionsvektoren sowie die Transformation in die entsprechenden bakteriellen Wirtszellen. Diese Arbeiten wurden von den jeweils verantwortlichen Kooperationspartnern durchgeführt und sollen an dieser Stelle nicht einzeln vorgestellt werden. Sie sind im Detail den entsprechenden Publikationen zu entnehmen [Alhapel *et al.* 2006; Bozarth 2005; Kress *et al.* 2009].

### 3.2 Proteinchemische Methoden

Sämtliche Arbeiten mit Mikroorganismen wurden gemäß Gentechnikgesetz (GenTG) vom 20.05.1990 durchgeführt und fielen unter die Sicherheitsstufe S1. Alle Arbeitsschritte wurden, sofern dies erforderlich war, unter sterilen Bedingungen ausgeführt, ohne dass dies bei der Beschreibung der jeweiligen Methoden explizit hervorgehoben wird. Dies gilt im Besonderen für die Herstellung von Medien und die Anzucht von Bakterienkulturen. Alle bei den Arbeiten mit Bakterien kontaminierten Lösungen und Gefäße wurden 60 min bei 121 °C und einem Druck von 2 bar sterilisiert.

Die Darstellung und die biochemische Charakterisierung der Enamidasen aus *Eubacterium barkeri* und *Bradyrhizobium japonicum*- wurden von Ashraf Alhapel bzw. Andrew Bozarth unter der Betreuung von Antonio Pierik durchgeführt und werden daher im Folgenden nicht gesondert besprochen. Die Vorgehensweise ist in der entsprechenden Publikation detailliert erläutert [Alhapel *et al.* 2006; Bozarth 2005]. Die enzymkinetischen Untersuchungen zur GcdA wurden von Daniela Bruegel übernommen [Kress *et al.* 2009]. Die heterologe Expression des *gcdA*-Gens sowie die Aufreinigung und Vorbereitung des Enzyms für die Kristallisation sind in Zusammenarbeit mit Daniela Bruegel und Iris Schall durchgeführt worden und werden im Folgenden beschrieben.

### 3.2.1 Heterologe Genexpression in *E. coli*

Die heterologe Expression des *gcdA*-Gens erfolgte im 1-6 L-Maßstab. Für einen typischen Expressionsansatz wurden zunächst jeweils 1 L LB-Medium in 5L-Schikanekolben autoklaviert. Dem Nährmedium wurden unmittelbar vor Beginn des Experiments 0.2% Glukose, 50 µM Biotin, Carbenicillin (50 µg/mL) and Chloramphenicol (35 µg/mL) hinzugefügt. Aus einer Übernachtskultur des entsprechenden Klons wurden anschließend jeweils 20 mL in die 5L-Kolben überführt und diese bei 37 °C und 150 U/min schütteln gelassen. Sobald eine Zelldichte von 0.5 bis 0.7 OD<sub>600</sub> erreicht war, wurde mit einer Endkonzentration von 0.4 µM Anhydrotetracyclin induziert und für 4 h bei 37 °C und 150 rpm inkubiert. Anschließend konnten die Zellen durch Zentrifugation (15 min, 6000 rpm, 4°C) in einer Kühlzentrifuge mit JA-10-Rotor geerntet werden. Die pelletierten Zellen wurden in einer möglichst geringen Menge Puffer W resuspendiert und in flüssigem Stickstoff schockgefroren. Die Lagerung erfolgte bei - 80 °C in der Tiefkühltruhe.

### 3.2.2 Zellaufschluss

Kleinere Ansätze bis 10 mL Zellsuspension wurden mit Hilfe des Ultraschallgerätes Nopplus HD 200 (*Bandelin electronic*) aufgeschlossen. Die Zellen wurden dazu mit 1 mM EDTA-, 0.2 µM PMSF- und 0.1 mg/mL Lysozym-Lösung versetzt und ca. 10 min bei RT inkubiert. Der Aufschluss erfolgte hiernach durch viermalige Wiederholung von 15 Ultraschallimpulsen (Programm 40, 72 D) unter Eiswasserkühlung.

Für Zellaufschlüsse in größerem Maßstab wurde ein so genannter Fluidizer (EmulsiFlex-C5, *Avestin*) verwendet. Bei dieser Methode wird auf die Zellsuspension zunächst ein hoher Druck ausgeübt. Beim Austritt der Suspension durch ein kleines Ventil erfolgt der Druckausgleich dann derart rasch, dass die Zellen durch die auftretenden Scherkräfte aufgebrochen werden. Je nach Ansatzgröße wurden 30-50 mL Zellsuspension auf die oben beschriebene Weise mit EDTA, PMSF und Lysozym versetzt und unter Eiskühlung aufgeschlossen. Der Vorgang wurde bis zu viermal wiederholt, um eine vollständige Lyse der Zellen zu gewährleisten.

Durch anschließende Zentrifugation (30 min, 13000 rpm, 4 °C) des Zellrohextrakts in einer Kühlzentrifuge mit JA-20-Rotor wurden unlösliche Zellbestandteile abgetrennt.



### 3.2.3 Chromatographische Proteinaufreinigung

Die chromatographische Aufreinigung der GcdA erfolgte durch Streptactin II-Affinitätschromatographie und anschließende Volumenausschlusschromatographie.

#### **Streptactin II-Affinitätschromatographie**

Nachfolgend soll lediglich auf die praktische Durchführung der Streptactin II-Affinitätsreinigung eingegangen werden. Ausführliche Informationen über die Funktionsweise der Methode sind der einschlägigen Literatur zu entnehmen [Schmidt & Skerra 2007].

Es wurde eine Strep-Tactin MacroPrep-Säule (IBA) mit einem Volumen von 5 mL verwendet. Zu Beginn der Aufreinigung wurde die Säule mit ca. 50 mL Puffer W äquilibriert. Das Zell-extrakt wurde partikelfrei gefiltert und mittels einer Perestaltikpumpe (GE Healthcare) mit einer Flussgeschwindigkeit von 0.5-1 mL pro Minute auf die Säule aufgetragen. Anschließend wurde die Chromatographiesäule an eines der zur Verfügung stehenden Chromatographiesysteme (vgl. 2.1) angeschlossen. Dabei wurde bei Flussraten von 0.5-2 mL/min gearbeitet. Um unspezifisch gebundene Kontaminationen von der Säule zu waschen, wurde im folgenden Schritt solange mit Puffer W gewaschen, bis keine Veränderung der Absorption bei 280 nm mehr zu beobachten war. Die Elution des Zielproteins erfolgte im Anschluss daran mit Puffer E. Zur Regeneration des Säulenmaterials wurde abschließend mit mehreren Säulenvolumen Puffer R gespült.

#### **Volumenausschlusschromatographie (Gelfiltration)**

Das Volumen der Elutionsfraktionen der Affinitätsaufreinigung wurde auf die in 3.2.4 beschriebene Weise reduziert, sterilfiltriert und zur weiteren Aufreinigung einer Volumenausschlusschromatographie unterzogen. Dazu wurde eine Superdex 200 10/300 GL-Gelfiltrationssäule (GE Healthcare) mit einem Volumen von 24 mL in Kombination mit einem ÄKTApurifier™ 100-System bei Flussraten von 0.3-0.5 mL/min verwendet. Abhängig von Gesamtvolumen und Konzentration der Proteinproben aus dem ersten Schritt der Aufreinigung wurden 0.1-0.5 mL auf die Säule aufgetragen. Als Laufpuffer wurde Puffer W verwendet. Zur Abschätzung des Molekulargewichts der GcdA wurde die Säule mit Hilfe der Proteine Thyroglobulin (669 kDa), Ferritin (440 kDa), Aldolase (158 kDa), Conalbumin (75 kDa) und Ovalbumin (43 kDa) kalibriert (Gelfiltrations-Kalibrierkit, GE Healthcare).

### 3.2.4 Konzentration der Proteinproben

Für die Konzentrierung der Proteinlösungen wurden Konzentratoren (Amicon-Ultra 4, *Milipore*) mit einem maximalen Fassungsvermögen von 4 mL bzw. 15 mL und einem Ausschlussvolumen von 30 kDa verwendet. Das Ankonzentrieren erfolgte bei 2000-4000 rpm in einer Kühlzentrifuge (Centrifuge 5910R, *Eppendorf*). Nach jeweils 1-5 min wurde die Zentrifugation unterbrochen und die Lösung mit einer Pipette durchmischt, um einer Ausfällung des Proteins vorzubeugen.

### 3.2.5 Entsalzung und Umpufferung

In einigen Fällen war es notwendig, die Zusammensetzung der Proteinlösung den Erfordernissen des jeweiligen Experiments (z.B. Konzentration und Art der Pufferkomponente, pH-Wert, Salzkonzentration) nachträglich anzupassen. Dies wurde entweder durch Dialyse in speziellen Dialyse-Kassetten (Slide-A-Lyzer® Mini Dialysis Units 10kDa, *Pierce*) oder mit Hilfe so genannter Entsalzungssäulen (HiTrap Desalting®-Säulen, *GE Healthcare*) durchgeführt.

### 3.2.6 Proteinkonzentrationsbestimmung

Die Proteinkonzentrationsbestimmung wurde mit Hilfe des 1985 von Smith entwickelten BCA-Tests durchgeführt [Smith *et al.* 1985]. Er beruht darauf, dass Peptidbindungen unter alkalischen Bedingungen  $\text{Cu}^{2+}$  zu  $\text{Cu}^+$  reduzieren. Die Menge an gebildetem  $\text{Cu}^+$  ist dabei proportional zur Proteinkonzentration.  $\text{Cu}^+$  bildet anschließend mit jeweils zwei Molekülen Bicinchoninsäure (engl.: bichoninic acid, BCA) einen violetten Komplex, der ein Absorptionsmaximum bei 562 nm aufweist. Der BCA-Test ist recht unempfindlich gegen Detergenzien wie Triton-X100 oder SDS, wird aber von komplexbildenden Reagenzien (z.B. EDTA) oder reduzierenden Substanzen wie beispielsweise Glucose, DTT oder Mercaptoethanol gestört. Außerdem ist die zu Grunde liegende Reduktion des Kupfers temperaturabhängig, weshalb für eine Vergleichbarkeit der Messung bei konstanter Temperatur gearbeitet werden muss.

Für den BCA-Test wurden je nach erwarteter Proteinkonzentration 0.5-50  $\mu\text{L}$  der Probe mit der entsprechenden Menge entionisiertem Wasser auf 50  $\mu\text{L}$  aufgefüllt und mit 1 mL frisch angesetzter BCA-Lösung (*Uptima*) versetzt. Die Mischung wurde für 30 min bei 37 °C inkubiert und anschließend die Absorption bei 562 nm gemessen. Bei jedem Test wurden parallel

zu den Proteinproben auch Standardproben mit bekannter Konzentration (0.05-1 mg/mL BSA) vermessen. Mit Hilfe dieser Standards wurde eine Kalibrationsgerade ermittelt und damit die Konzentration berechnet.

### 3.2.7 SDS-Polyacrylamidgelelektrophorese (SDS-PAGE)

Die Technik der SDS-Polyacrylamidgelelektrophorese (SDS-PAGE) wurde zur qualitativen Proteinanalyse und zur Reinheitskontrolle der aufgereinigten Proteine genutzt. Die Methode beruht auf den denaturierenden Eigenschaften des Natriumdodecylsulfats (engl.: sodium dodecylsulfate, SDS). Die Dodecylsulfat-Anionen binden unter Bildung eines negativ geladenen SDS-Protein-Komplexes fest an das Protein, wobei die Ladung des Proteins „maskiert“ wird. Das Polyacrylamidgel stellt ein molekulares Sieb dar. Wird ein elektrisches Feld angelegt, findet darin eine Trennung der aufgetragenen Proteine nach ihren molaren Massen statt. Die Geschwindigkeit, mit der der jeweilige SDS-Protein-Komplex die Acrylamidmatrix passiert, ist proportional zum Logarithmus der korrespondierenden Proteingröße und abhängig von der angelegten Spannung. Ein typisches SDS-Gel besteht aus zwei Teil-Gelen mit verschiedenen Porengrößen und pH-Werten: dem weitporigen Sammelgel (pH 6.8) und dem sich anschließenden engporigen Trenngel (pH 8.8). Auf Grund der unterschiedlichen pH-Werte liegt die Pufferkomponente Glycin im Sammelgel zwitterionisch, im Trenngel einfach negativ geladen vor. Im elektrischen Feld wandern nur die geladenen Chloridionen durch das Sammelgel, während das zwitterionische Glycin zurückbleibt. Beim Eintritt ins Trenngel werden Feldstärke und Porengröße – und damit die Mobilität der SDS-Protein-Komplexe – abrupt reduziert. Dies führt dazu, dass sich die Komplexe an der Grenzlinie in einer schmalen Bande sammeln. Da im Trenngel das einfach negativ geladene Glycin-Anion vorliegt, wird der Mangel an beweglichen Ionen aufgehoben, und es ergibt sich eine lediglich von ihrer Größe abhängende Wanderungsgeschwindigkeit der Proteine.

In dieser Arbeit kam ausschließlich die SDS-PAGE nach Laemmli [Laemmli 1970] zur Anwendung. Die Elektrophorese wurde in kleinen vertikalen Gelkammern („mighty small“, *Hofer Scientific*) bei Stromstärken zwischen 15 und 25 mA durchgeführt. Am häufigsten wurde eine Kombination aus 4 %igem Sammelgel und 12 bzw. 15 %igem Trenngel benutzt. Die genaue Zusammensetzung ist der nachfolgenden Tabelle zu entnehmen, wobei die angegebene Menge für jeweils zwei Gele ausgelegt ist. Die Proteinproben wurden vor dem Auftragen auf das Gel mit Probenpuffer versetzt und 5 min bei 95 °C im Heizblock denatu-

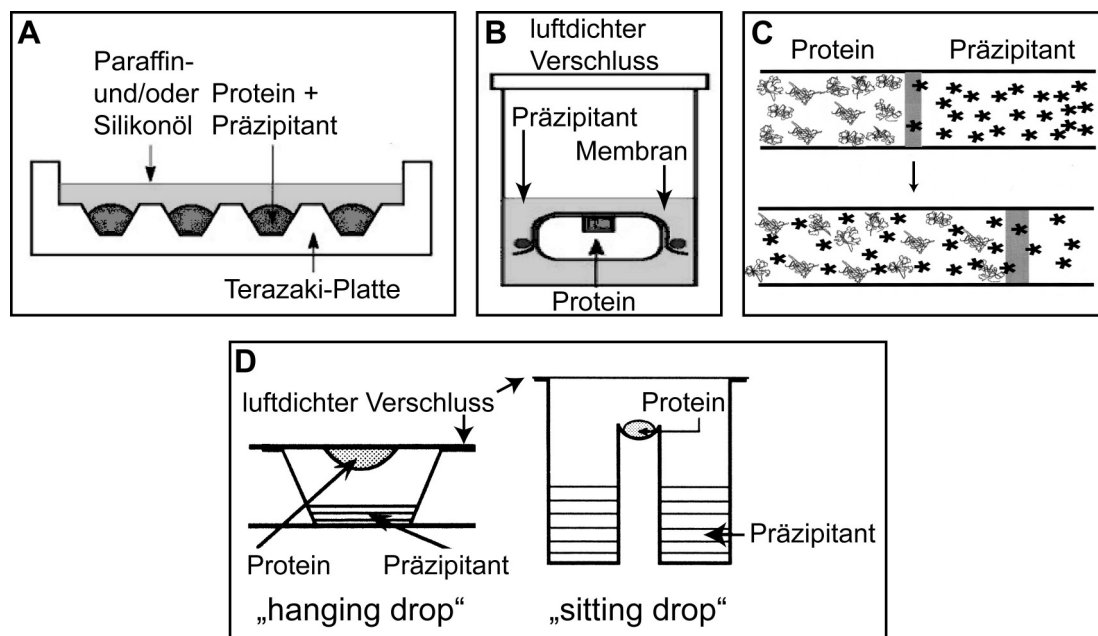
riert. Für die Zuordnung des Molekulargewichts der zu beobachtenden Banden wurde stets ein Größenstandard mit aufgetragen (peqGOLD Protein-Marker I, *PEQLAB*). Im Anschluss an die Elektrophorese wurden die Gele für mindestens 30 min in Coomassie Blue-Färbelösung [Bennett & Scott 1971] und danach in kochendem Entfärber entfärbt.

Laemmli-Gele	Trenngel 12 %	Trenngel 15 %	Sammelgel 4 %
Acrylamidlösung	4 mL	5 mL	1.33 mL
dest. H <sub>2</sub> O	2.53 mL	1.53 mL	6.57 mL
10 % (w/v) SDS	100 µL	100 µL	100 µL
Trenngelpuffer	3.33 mL	3.33 mL	-
Sammelgelpuffer	-	-	2.0 mL
TEMED	10 µL	10 µL	10 µL
10 % (w/v) APS	50 µL	50 µL	100 µL

### 3.3 Proteinkristallisation

Eine notwendige Bedingung für die röntgenographische Strukturbestimmung von Proteinen ist die Verfügbarkeit gut streuender Kristalle. Welches die geeigneten Bedingungen für die Kristallisation des jeweiligen Proteins sind, lässt sich nur durch empirische Methoden zweifelsfrei feststellen. Diese Methoden beinhalten die systematische Variation zahlreicher Faktoren, die für die Proteinkristallisation von Bedeutung sein können: Temperatur, pH-Wert, Pufferart, Konzentration und Art des Fällungsmittels, Ionenstärke sowie die Konzentration und Reinheit des Proteins.

Alle Methoden der Proteinkristallisation beruhen im Prinzip darauf, dass durch die schonende Erhöhung der Proteinkonzentration und/oder der Konzentration eines oder mehrerer so genannter Fällungsreagenzien (Präzipitantien) eine übersättigte Proteinlösung (Mutterlauge) entsteht, aus der sich das Zielprotein in Form wohlgeordneter Kristalle abscheidet. Je nach experimentellem Aufbau und den zu Grunde liegenden physikalischen Effekten werden im Wesentlichen vier Methoden der Kristallisation unterschieden (Abbildung 3-1) [Ducruix & Giegé 1999; McPherson 1999].



**Abbildung 3-1 Methoden der Proteinkristallographie** (A) Batchmethode (B) Mikrodialyse am Beispiel eines „Dialyseknopfs“ (C) free interface-Diffusion (D) *hanging* und *sitting drop*-Dampfdiffusion

### **Batchmethode**

Die Batchmethode ist die älteste Methode der Proteinkristallisation. Dabei werden Protein- und Präzipitantlösung unter einer mehr oder weniger luftdichten Schicht aus Paraffin und/oder Silikonöl vermischt. Konzentrationsänderungen sind hier nur durch das Ausfallen einer der Komponenten zu erwarten, z.B. durch teilweises Ausfallen bzw. die Kristallisation des Proteins. Die Methode wird heutzutage vor allem für die Kristalloptimierung verwendet.

### **Mikrodialyse**

Hier wird die Konzentrationsänderung durch Dialyse erreicht. Diese Methode kann beispielsweise Erfolg versprechend sein, wenn das Zielprotein nur unter Hochsalzbedingungen löslich ist. Eine behutsame Dialyse gegen geringere Salzkonzentrationen kann dann dazu führen, dass es sich in Form von Kristallen abscheidet, anstatt amorph auszufallen.

### **Free interface-Diffusion (dt.: Diffusion an freien Grenzflächen)**

Diese Technik wurde entscheidend von Gracia-Ruiz geprägt, der sie zur so genannten counter diffusion-Methode (dt. „Gegendiffusionsmethode“) weiterentwickelt hat. Hierbei vermischen sich Protein- und Präzipitantlösung nach und nach innerhalb einer langgestreckten Kapillare. Entlang der Kapillare entsteht so ein Protein-Präzipitant-Gradient. Auf diese Weise kann eine ganze Reihe von Protein-zu-Präzipitant-Verhältnissen gleichzeitig getestet und mit Hilfe spezieller mikroskopischer Techniken begutachtet werden [Garcia-Ruiz *et al.* 2002].

### **Dampfdiffusion**

Die Dampfdiffusion ist die bei Weitem am häufigsten verwendete Methode der Proteinkristallisation. Sie beruht auf einem einfachen Prinzip: In einem luftdicht verschlossenen Reaktionsgefäß stehen sich ein Tropfen aus Protein- und Präzipitantlösung und ein deutlich größeres Volumen der unverdünnten Präzipitantlösung gegenüber; die Zusammensetzung des Tropfens nähert sich durch Dampfdiffusion der annähernd konstanten Konzentration des Reservoirs an. Je nach Position des Tropfens im Reaktionsgefäß spricht man von der Technik des „sitzenden- bzw. hängenden Tropfens“ (engl.: *sitting* bzw. *hanging drop*) [Davies & Segal 1971]. Letztere Technik wird oft in der Kristalloptimierung eingesetzt, da sie das Wachstum größerer Kristalle begünstigt und gleichzeitig verhindert, dass diese an der Oberfläche des Reaktionsgefäßes „festwachsen“.

Im Rahmen der hier vorgestellten Arbeit wurde ausschließlich die Dampfdiffusionsmethode eingesetzt. Um unerwünschte Aggregation zu unterbinden, wurden alle Proteinproben und eingesetzten Lösungen vor den Kristallisationsexperimenten partikelfrei filtriert.

### 3.3.1 Automatisiertes Kristallisationsscreening

Im ersten Schritt der Gewinnung qualitativ hochwertiger Proteinkristalle für die röntgenographische Strukturbestimmung wird versucht, die physikochemischen Bedingungen (Protein- und Fällungsmittelkonzentration, pH-Wert, Temperatur usw.) einzugrenzen, unter denen das Zielprotein kristallisiert. Dieser Vorgang wird als Screening bezeichnet.

Im Rahmen dieser Arbeit wurde für das Screening der Kristallisationsroboter Cartesian Microsys<sup>TM</sup> SQ4000 (*Genomic Solutions*) in Kombination mit vorgefertigten Kristallisationslösungen (Screens) der Firmen *Sigma-Aldrich* und *Qiagen* verwendet. Es handelt sich dabei um Sammlungen besonders erfolgreicher Kristallisationsbedingungen und systematischer Variationen von diesen. Das Prinzip dieser so genannten *sparse matrix*-Screens geht auf die Arbeiten von Jancarik & Kim zurück [Jancarik & Kim 1991]. Die Kristallisation erfolgte in speziellen Reaktionsgefäßen der Firmen *Greiner Bio-one* und *Hampton Research*, die als Kristallisationsplatten bezeichnet werden. Sie fassen bis zu 96 verschiedene Kristallisationsbedingungen im 8x12-Format und je nach Ausführung bis zu drei Vertiefungen je Reservoir für die *sitting drop*-Kristallisation.

Für ein typisches Screening wurden zunächst jeweils 80-100 µL der Kristallisationslösungen in die Reservoirs pipettiert. Die Protein-Präzipitant-Tropfen (300-500 nL Gesamtvolumen) wurden im Anschluss daran vom Kristallisationsroboter erzeugt. Nachdem die Kristallisationsplatten mit einer speziellen Folie (VIEWseal<sup>TM</sup> Klebefolien, *Greiner Bio-one*) luftdicht verschlossen worden waren, erfolgte die Inkubation der Ansätze bei 18 °C in speziellen Inkubationsschränken oder bei 4-6 °C im Kühlraum. Direkt nach Fertigstellung der Platten wurden die einzelnen Tropfen überprüft und eine gegebenenfalls eingetretene Trübung der Lösung oder die Präzipitation des Proteins protokolliert. Dazu wurden die Kristallisationsansätze unter einem Mikroskop durch einen Polfilter hindurch betrachtet oder mit Hilfe des automatisierten Dokumentationssystems Crystal Pro HT 110 (*TriTek Corporation*) begutachtet. Die Ansätze wurden nach 2-3 Tagen erneut und danach etwa einmal pro Woche auf die gleiche Weise kontrolliert.

### 3.3.2 Kristallisationsoptimierung

Oft sind die Kristalle, die im Rahmen eines ersten Screenings erhalten wurden, noch zu klein oder von unzureichender Qualität für die röntgenographische Vermessung. Daher ist es in den meisten Fällen notwendig, eine Optimierung der entsprechenden Kristallisationsbedingungen vorzunehmen. Dabei werden in systematischer Weise Parameter wie die Zusammensetzung und der pH-Wert der Kristallisationslösungen, die Proteinkonzentration, die Inkubationstemperatur und die Tropfengröße variiert.

In dieser Arbeit wurde die Kristalloptimierung zumeist von Hand im *hanging drop*-Verfahren durchgeführt. Die hierzu ursprünglich verwendeten VDX-Platten der Firma *Hampton Research* wurden später aus praktischen Gründen vollständig durch wiederverschließbare *EasyXtal Tools* (*Qiagen*) ersetzt. Letztere besitzen ebenfalls 24 Reservoirs im 4x6 Format und werden mit einem Schraubverschluss luftdicht verschlossen. Die Wiederverschließbarkeit der *EasyXtal Tools* erleichtert die nachträgliche Manipulation der Kristalle erheblich. Für ein typisches Optimierungsexperiment wurden 0.5-4  $\mu\text{L}$  der Proteinprobe mit dem gleichen Volumen der jeweiligen Kristallisationslösung auf dem Schraubverschluss vermischt und dieser anschließend auf das mit 0.75-1 mL Kristallisationslösung befüllte Reservoir aufgeschraubt. Um die optimalen Bedingungen für das Kristallwachstum zu finden, wurden Protein- und Präzipitantkonzentration sowie Art und pH-Wert der Pufferkomponente variiert. In einigen Fällen wurde durch Zugabe von Glycerin versucht, den Prozess der Kristallisation zu verlangsamen bzw. unerwünschte Proteinaggregation zu verhindern und damit die Qualität der Kristalle zu verbessern. Die Herstellung der Bedingungen erfolgte je nach Umfang des jeweiligen Experiments manuell oder automatisiert mit Hilfe des Liquid handling-Systems *Lissy* (*Zinsser Analytic*). Darüber hinaus kamen so genannte Additiv-Screens verschiedener Hersteller zum Einsatz (*Sigma Aldrich*, *Hampton Research*), die eine ganze Reihe unterschiedlicher Verbindungen enthalten, die sich günstig auf die Kristallisation auswirken können. Zu Beginn einer Optimierung wurden die Ansätze meist in doppelter Ausführung angefertigt und anschließend sowohl bei 18 °C als auch bei Kühlraumtemperatur von 4-6 °C inkubiert.

### 3.3.3 Kokristallisation und *soaking*-Experimente

Handelt es sich bei dem betrachteten Zielprotein um ein Enzym, ist es möglich, mit Hilfe des Strukturmodells Informationen über die katalysierte Reaktion zu erhalten. Daher ist es wünschenswert, die strukturelle Charakterisierung in Gegenwart von Substraten (Produkt(e)



bzw. Edukt(e)), Substratanaloga oder Inhibitoren bzw. Aktivatoren des Enzyms durchzuführen. Oft sind auch prosthetische Gruppen wie z.B. Metalle oder Kofaktoren – etwa Biotin oder FAD – für die Funktion des Proteins von besonderer Bedeutung. Da diese auch als *Liganden* bezeichneten Gruppen jedoch in vielen Fällen nicht-kovalent oder nur mit unzureichender Affinität an das Zielprotein binden, gehen sie im Rahmen der Aufreinigung des Proteins oft „verloren“. Außerdem ist bei der Darstellung durch heterologe Expression nicht immer gewährleistet, dass der Expressionswirt die entsprechenden Verbindungen (in ausreichenden Mengen) zur Verfügung stellen kann. In solchen Fällen ist es möglich, die Liganden dem Protein vor der Kristallisation hinzuzufügen bzw. nachträglich in den Kristall zu inkorporieren. Diese Methoden werden als *Kokristallisation* bzw. *soaking* (dt.: einweichen) bezeichnet.

### **Soaking**

Im Vergleich zu Kleinmolekülstrukturen weisen Proteinkristalle einen beträchtlichen Lösungsmittelanteil auf (oft weit über 50%) und sind durch große Abstände im Kristallgitter und solvensgefüllte Kanäle in den einzelnen Proteinmolekülen gekennzeichnet. Es ist daher prinzipiell möglich, kleinere Moleküle nachträglich in die Kristalle einzuführen. Dazu werden die Verbindungen entweder direkt der Mutterlauge hinzugefügt, welche die Kristalle enthält, oder aber die Kristalle werden manuell in eine Lösung des entsprechenden Liganden überführt. Die Zeitdauer eines *soaking*-Experiments richtet sich nach der Stabilität der Kristalle bzw. des Liganden und liegt üblicherweise bei einer bis mehreren Minuten, kann sich aber in einigen Fällen auch über den Zeitraum von mehreren Stunden erstrecken.

### **Kokristallisation**

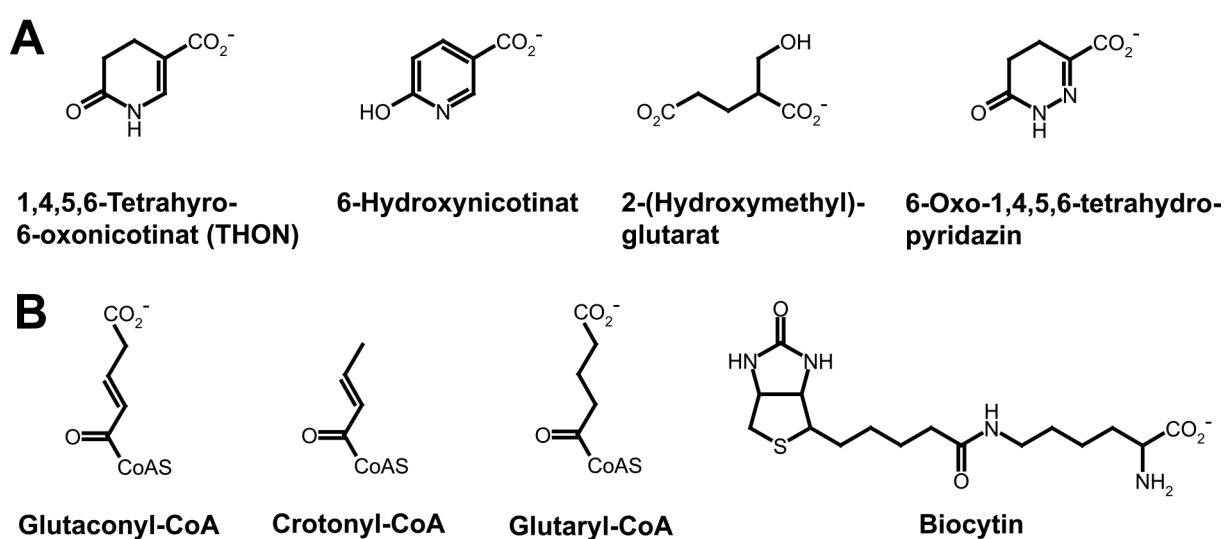
Der nachträgliche Einbau von Molekülen in den Kristall kann durch die Wechselwirkung des Liganden mit dem Protein teilweise zu beträchtlichen Veränderungen im Kristallgitter führen. In diesen Fällen besteht die Gefahr, den Kristall durch die Anwendung der *soaking*-Methode stark zu beschädigen und somit unbrauchbar für die strukturelle Charakterisierung zu machen. In Kokristallisationsexperimenten werden daher die entsprechenden Verbindungen der Proteinlösung schon vor der Kristallisation hinzugefügt. Allerdings kann es passieren, dass es unter einer zuvor ermittelten Bedingung auf Grund der Wechselwirkungen zwischen Substrat und Protein nicht zu Ausbildung von Kristallen kommt oder diese nicht von ausreichender Qualität für röntgenographische Experimente sind. In solchen Fällen ist

ein erneutes Screening sinnvoll, vorausgesetzt, der entsprechende Ligand ist in ausreichenden Mengen verfügbar.

Im Rahmen der vorliegenden Arbeit wurden Kokristallisations- bzw. *soaking*-Experimente sowohl für Enamidase als auch für GcdA durchgeführt. Die Strukturformeln der hierfür eingesetzten Chemikalien sind in Abbildung 3-2 gezeigt. Für Kokristallisations- und *soaking*-Experimente mit Enamidase wurden neben dem Substrat 1,4,5,6-Tetrahydro-6-oxonicotinat (THON) und der strukturell ähnlichen Verbindung 6-Oxo-1,4,5,6-tetrahydro-pyridazin der THON-Vorläufer 6-Hydroxynicotinat und 2-(Hydroxymethyl)-glutarat verwendet. Bei letzterer Verbindung handelt es sich um das Produkt des in der Nicotinat-Fermentation von *E. barkeri* auf Enamidase folgenden Enzyms 2-(Hydroxymethyl)-glutarat-Dehydrogenase.

Im Falle von GcdA wurde das Substrat Glutaconyl-CoA, das Produkt Crotonyl-CoA und das Substratanalogon Glutaryl-CoA eingesetzt. Statt dem Kosubstrat Biotin wurde die deutlich besser wasserlösliche Verbindung Biocytin verwendet.

Für das Enamidase-Projekt wurden die Kokristallisationsexperimente ausschließlich ausgehend von bereits ermittelten Kristallisationsbedingungen durchgeführt. Bei der GcdA erfolgte zusätzlich zu dieser Vorgehensweise ein erneutes Screening in Gegenwart des/der jeweiligen Liganden. Die Menge der eingesetzten Liganden richtete sich nach Löslichkeit der jeweiligen Verbindung und den enzymkinetischen Eigenschaften des Zielproteins.



**Abbildung 3-2 In Kokristallisations- und *soaking*-Experimenten eingesetzte Verbindungen**  
(A) Enamidase (B) GcdA

## 3.4 Proteinstrukturbestimmung

### 3.4.1 Kristalle und Symmetrie

Ein Kristall ist ein Festkörper, dessen atomare Bestandteile in Form eines gleich bleibenden Strukturmusters angeordnet sind, das sich periodisch in allen drei Raumrichtungen wiederholt. Die kleinste „Baueinheit“ des auf diese Weise entstehenden *Translations-* bzw. *Kristallgitters* stellt die *Elementar-* bzw. *Einheitszelle* dar, die durch die *Gitterparameter*  $a, b, c$  (die Beträge der *Basisvektoren* des Kristallgitters) und die drei zwischen den Basisvektoren aufgespannten Winkel  $\alpha, \beta, \gamma$  charakterisiert wird (Abbildung 3-3). Die Lage eines Atoms innerhalb des Gitters wird mit Hilfe der *Atomkoordinaten*  $x, y, z$  angegeben, die Bruchteile der Gitterkonstanten darstellen ( $ax, ay, az$ ). Die meisten Kristalle besitzen Symmetrieeigenschaften, welche die Symmetrie des Translationsgitters beeinflussen. Auf Grund der Restriktionen, die sich daraus für die Gitterparameter ergeben, können insgesamt 7 verschiedene *Kristallsysteme* unterschieden werden (Tabelle 3-1). Beschreibt die Elementarzelle die kleinste repräsentative Volumeneinheit des Kristallgitters, so spricht man von einer *primitive Zelle* ( $P$ ). Sie enthält einen *Translations-* oder *Gitterpunkt*, der sich zu gleichen Teilen auf die Ecken der Zelle verteilt. Die Elementarzelle wird jedoch stets so definiert, dass sie mit dem Kristallsystem der höchstmöglichen Symmetrie vereinbar ist. Es ist daher oft notwendig, größere Zellen zu wählen, die zusätzliche Gitterpunkte in sich einschließen. Neben der primitiven Zelle unterscheidet man: *flächenzentrierte Zellen* ( $C$ ), in denen ein zusätzlicher Translationspunkt auf der Mitte einer der Flächen sitzt; *allseitig-flächenzentrierte Zellen* ( $F$ ); *raum-* bzw. *innenzentrierte Zellen* ( $I$ ), bei denen ein zusätzlicher Translationspunkt in der Zellmitte sitzt. Insgesamt ergeben sich so in Kombination mit den 7 Kristallklassen 14 so genannte *Bravais-Gitter* (Abbildung 3-4).

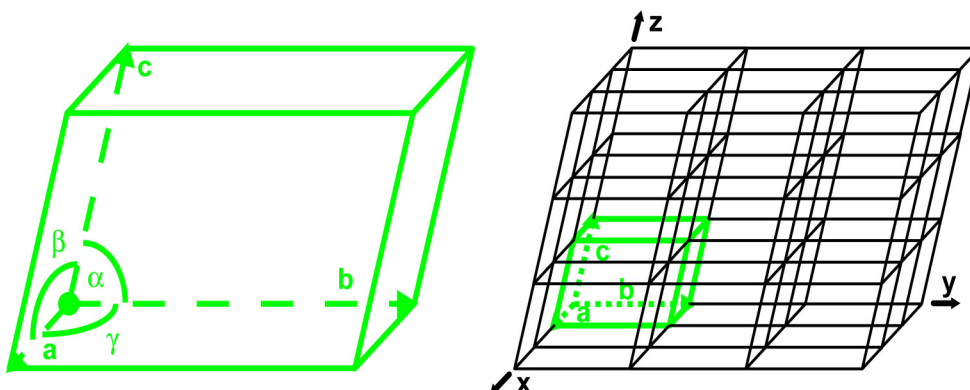


Abbildung 3-3 Einheitszelle und das resultierende Kristallgitter

Der kleinste nicht-reduzierbare Baustein, mit dessen Hilfe man durch Symmetrieeoperation die komplette Einheitszelle und damit das gesamte Kristallgitter beschreiben kann, wird als *asymmetrische Einheit* bezeichnet. Durch die Kombination einfacher (*Rotation*, *Spiegelung* und *Inversion*) und gekoppelter Symmetrieelemente ergeben sich 32 *Punktgruppen*, die mit den Gittertypen kompatibel sind. Die Anwendung der Translationssymmetrie der 14 Bravais-Gitter führt schließlich zu den 230 *Raumgruppen*, welche Art und Lage der Symmetrieelemente in einer Kristallstruktur vollständig beschreiben. Da in Proteinkristallen auf Grund ihrer Chiralität nur *Translations*-, *Rotations*- und *Schraubenachsen* möglich sind, reduziert sich bei ihnen die Anzahl der möglichen Raumgruppen auf 65 (Tabelle 3-1).

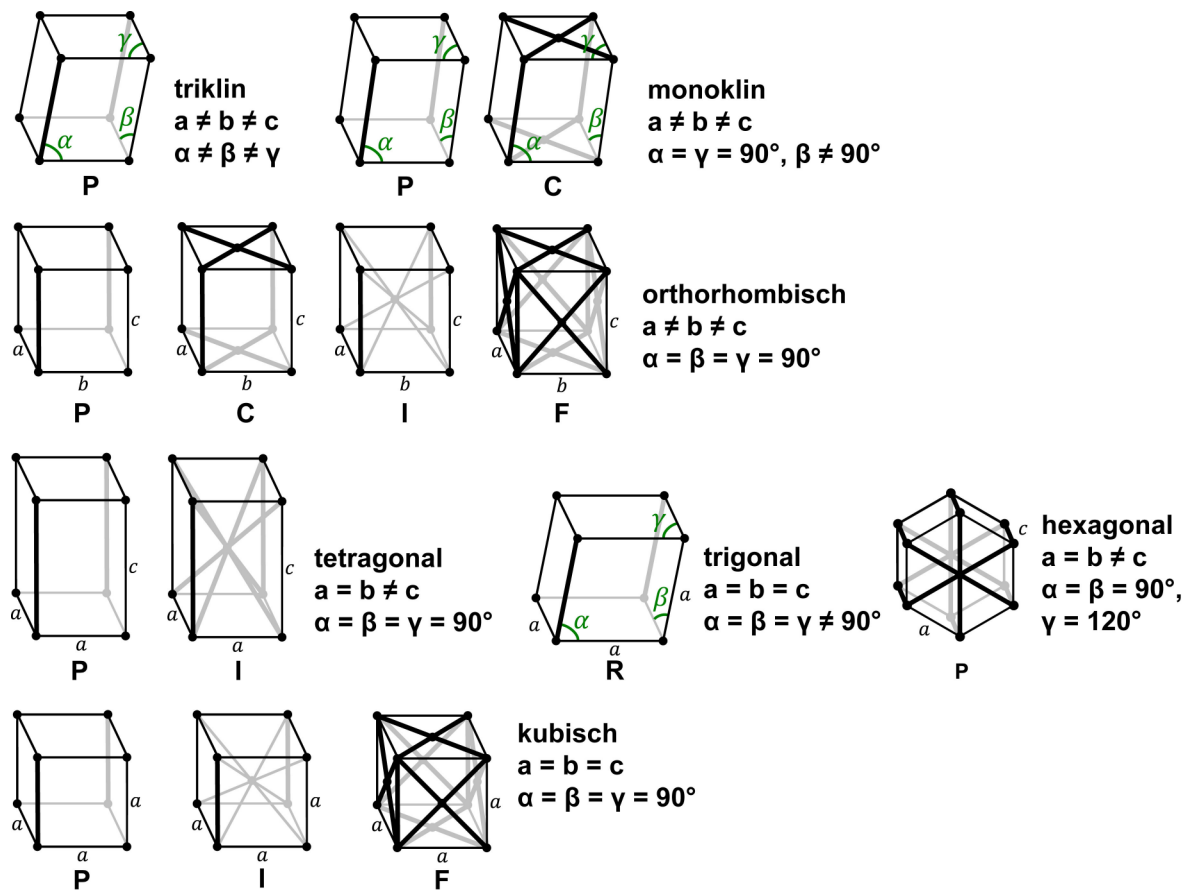


Abbildung 3-4 Die 14 Bravais-Gitter

Kristallsystem	Gittertyp	minimale Symmetrie der EZ	Gitterparameter	Punktgruppen	erlaubte Raumgruppen
triklin	P	keine	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$	1	P1
monoklin	P	einzelne zweizählige Achse	$a \neq b \neq c$ $\alpha = \gamma = 90^\circ$ $\beta \neq 90^\circ$	2	P2, P2 <sub>1</sub> , C2
orthorhombisch	P C I F	drei aufeinander senkrecht stehende zweizählige Achsen	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	222	P222, P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> , P222 <sub>1</sub> , P2 <sub>1</sub> 2 <sub>1</sub> 2 C222, C222 <sub>1</sub> I222, I2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub> F222
tetragonal	P I	einzelne vierzählige Achse	$a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	4  422	P4, P4 <sub>1</sub> , P4 <sub>2</sub> , P4 <sub>3</sub> I4, I4 <sub>1</sub>  P422, P4 <sub>2</sub> 2, P4 <sub>1</sub> 22, P4 <sub>1</sub> 2 <sub>1</sub> 2, P4 <sub>2</sub> 22, P4 <sub>2</sub> 2 <sub>1</sub> 2, P4 <sub>3</sub> 22, P4 <sub>3</sub> 2 <sub>1</sub> 2 I422, I4 <sub>1</sub> 22
trigonal	R P	einzelne dreizählige Achse	$a = b = c$ $\alpha = \beta = \gamma \neq 90^\circ$	3  32	R3 P3, P3 <sub>1</sub> , P3 <sub>2</sub>  R32 P321, P312 P3 <sub>1</sub> 12, P3 <sub>1</sub> 21, P3 <sub>2</sub> 12, P3 <sub>2</sub> 21
hexagonal	P	einzelne sechszählige Achse	$a = b \neq c$ $\alpha = \beta = 90^\circ$ $\gamma = 120^\circ$	6  622	P6, P6 <sub>1</sub> , P6 <sub>5</sub> , P6 <sub>3</sub> , P6 <sub>2</sub> , P6 <sub>4</sub>  P622, P6 <sub>1</sub> 22, P6 <sub>5</sub> 22, P6 <sub>2</sub> 22, P6 <sub>4</sub> 22, P6 <sub>3</sub> 22
kubisch	P I F	dreizählige Achse entlang der Würfel-diagonale	$a = b = c$ $\alpha = \beta = \gamma = 90^\circ$	23  432	P23, P2 <sub>1</sub> 3 I23, I2 <sub>1</sub> 3 F23 P432, P4 <sub>2</sub> 32, P4 <sub>3</sub> 32, P4 <sub>1</sub> 32 I432, I4 <sub>1</sub> 32 F432, F4 <sub>1</sub> 32

Tabelle 3-1 Einheitszellen und erlaubte Raumgruppen für biologische Makromoleküle

Es ist möglich, dass die Anzahl asymmetrischer Einheiten in der Elementarzelle kleiner ist als die Gesamtanzahl der Moleküle. In diesem Fall besitzt die asymmetrische Einheit keine oder lokale, so genannte *nichtkristallographische Symmetrie* (engl.: *non-crystallographic symmetry*, NCS), welche von der Raumgruppe unabhängig zu betrachten ist. Im Falle von Proteinkristallen handelt es sich dabei häufig um eine geschlossene Punktgruppensymmetrie (engl.: *proper NCS*), bei der die chiralen Moleküle durch eine n-zählige Rotation ineinander überführt werden. NCS kann jedoch ebenso aus einer Kombination von Rotation und Translation bestehen. Man spricht dann von offener NCS (engl.: *improper NCS*).

Da NCS für den Gang einer Strukturlösung von großer Bedeutung sein kann, wird in der Praxis versucht, nach der Bestimmung von Raumgruppe und Gitterparametern die Anzahl  $Z$  der Moleküle in der asymmetrischen Einheit abzuschätzen. Dies erfolgt am einfachsten mit der 1968 von Matthews vorgeschlagenen Methode [Matthews 1968]: Er analysierte den Lösungsmittelgehalt 116 verschiedener Proteinkristallformen und fand heraus, dass das Verhältnis „Einheitszell-Volumen“/„Molekulargewicht des Proteins“ im Mittel  $2.15 \text{ \AA}^3/\text{Da}$  beträgt (Matthews-Koeffizient,  $V_M$ ). Die Zusammensetzung der asymmetrischen Einheit lässt sich somit abschätzen, indem man diesen Quotienten für verschiedene Werte von  $Z$  berechnet, mit  $V_M$  vergleicht und die Symmetrie der Raumgruppe mitberücksichtigt.

Den Zusammenhang zwischen asymmetrischer Einheit, Elementarzelle und Translationsgitter hat McPherson in seiner Monographie „Introduction to macromolecular crystallography“ anschaulich illustriert (vgl. Abbildung 3-5).

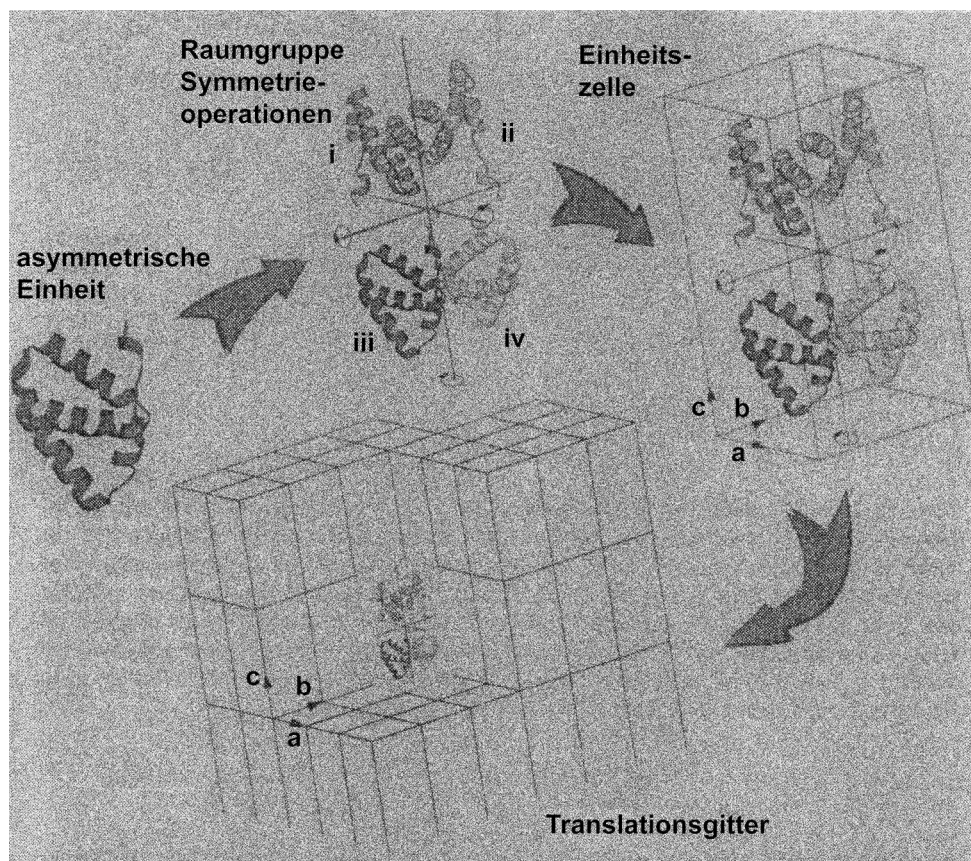


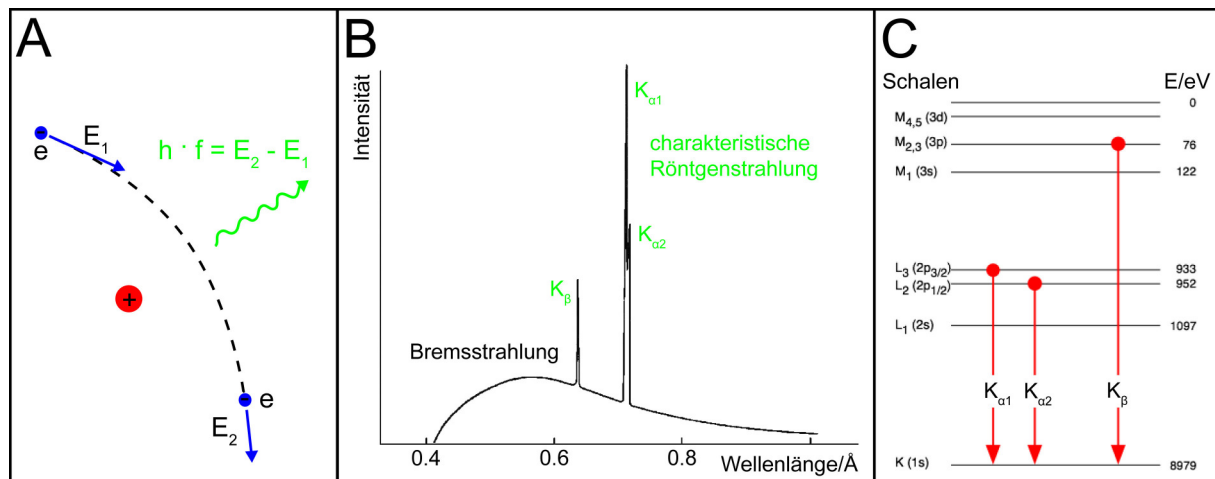
Abbildung 3-5 Schematischer Aufbau eines Proteinkristalls durch die symmetrische Wiederholung einfacher Grundelemente [McPherson 2003]

### 3.4.2 Erzeugung von Röntgenstrahlen

Als *Röntgenstrahlung* bezeichnet man elektromagnetische Wellen einer Photonenenergie zwischen ungefähr 100 eV und 250 keV und von Wellenlängen zwischen  $10^{-8}$  und  $10^{-12}$  m (0.001 bis 100 Å). Der Wellenlängenbereich liegt in der Größenordnung der atomaren Abstände innerhalb von Kristallen (Größenordnung 1 Å) und wird daher für Beugungsexperimente in der Kristallographie eingesetzt. Die Erzeugung von Röntgenstrahlung für die Strukturanalyse beruht im Wesentlichen auf zwei Effekten: der so genannten *Bremsstrahlung* und der *charakteristischen Röntgenstrahlung*.

Bremsstrahlung (*weiße Röntgenstrahlung*) entsteht, wenn ein geladenes Teilchen, zum Beispiel ein Elektron, beschleunigt bzw. abgebremst wird (Abbildung 3-6A). Im Falle einer einfachen Röntgenröhre tritt dieses Phänomen durch das Abbremsen energiereicher Elektronen beim Eintritt in das verwendete Anodenmaterial (zumeist Mo oder Cu) auf. Die charakteristische Röntgenstrahlung ist dem kontinuierlichen Spektrum der Bremsstrahlung überlagert (Abbildung 3-6B). Sie entsteht, wenn ein Rumpfelektron der Metallanode (z.B. aus der K-Schale) durch Kollision mit dem Elektronenstrahl unter Ionisierung des Atoms herausgeschlagen wird. Dieser instabile Zustand relaxiert sofort durch den Sprung eines Elektrons aus einer höheren Schale (z.B. der L-Schale) in die Lücke der K-Schale. Dabei wird Röntgenstrahlung scharf definierter Wellenlänge emittiert, die sich aus der Energiedifferenz beider Niveaus ergibt. Für den Übergang von der L- zur K-Schale (Auswahlregel  $\Delta l = \pm 1$ ) wird ein Dublett mit eng benachbarten Wellenlängen emittiert: die  $K_{\alpha 1}$ - und  $K_{\alpha 2}$ -Strahlung (Abbildung 3-6C). Da für die meisten Beugungsexperimente monochromatische Strahlung benötigt wird, benutzt man die besonders intensive charakteristische  $K_{\alpha}$ -Strahlung und versucht, störende Strahlung anderer Wellenlängen möglichst zu eliminieren.

In der Röntgenstrukturanalyse von biologischen Makromolekülen werden häufig so genannte *Drehanoden-Generatoren* als Röntgenquellen eingesetzt. Bei dieser Weiterentwicklung der konventionellen wassergekühlten Vakuumröntgenröhre wird ein schnell rotierender Metallzylinder als Anode verwendet, wodurch die Wärme besser abgeführt werden kann. Dadurch sind höhere Leistungen und somit höhere Strahlungsintensitäten möglich.



**Abbildung 3-6 Zur Erzeugung von Röntgenstrahlung** (A) Entstehung von Bremsstrahlung durch Ablenkung eines energiereichen Elektrons im Coulombfeld eines Atomkerns (B) Spektrum einer Molybdän-Röntgenröhre [Massa 2007] (C) Die ersten drei K-Linien von Kupfer

Kristalle von biologischen Makromolekülen, wie beispielsweise Proteinen, enthalten deutlich weniger Moleküle als gleich große Kristalle kleinerer Moleküle und beugen Röntgenstrahlung daher nur vergleichsweise schwach. Außerdem bestehen Proteine größtenteils aus den leichten Elementen Kohlenstoff, Stickstoff und Sauerstoff, die Röntgenstrahlung aufgrund der geringen Anzahl an Elektronen nur schwach streuen. Darüber hinaus ist es häufig schwierig, Proteinkristalle von ausreichender Größe und Qualität reproduzierbar herzustellen. Daher gewinnt die Verwendung der so genannten Synchrotronstrahlung immer mehr an Bedeutung. Sie verfügt über einige wesentliche Vorteile:

- hohe Intensität bei geringer Divergenz
- durchstimmbare Wellenlänge
- hoher Polarisationsgrad

Synchrotronstrahlung ist eine besondere Form der bereits erwähnten Bremsstrahlung. Sie wird zumeist in so genannten Elektronen-Speicherringen erzeugt, in denen beschleunigte Elektronen durch Magnetfelder auf einer Kreisbahn gehalten werden, wobei sie tangential zu ihrer Bewegungsrichtung ein breites Spektrum elektromagnetischer Strahlung emittieren. Man unterscheidet Strahlungsquellen der ersten, zweiten, dritten und vierten Generation. Ihr zentrales Unterscheidungsmerkmal ist die Brillanz der emittierten Strahlung. Sie ist definiert als die Zahl der Photonen pro Fläche, Raumwinkel und Zeit innerhalb eines schmalen Wellenlängenbereichs.

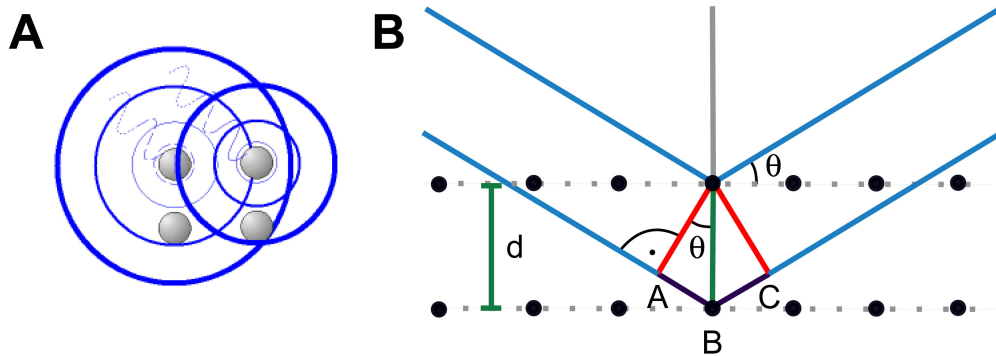


Zunächst wurde Synchrotronstrahlung „parasitär“ an Teilchenbeschleunigern der Teilchenphysik gewonnen, die als Synchrotronstrahlungsquellen erster Generation bezeichnet werden. Ab der zweiten Generation wurden Speicherringe gebaut, die allein der Erzeugung der Synchrotronstrahlung dienen. Die Elektronen werden dabei in speziellen Magnetstrukturen, den *Dipolmagneten* und *Wiggeln*, beschleunigt. Ein Wiggler ist eine lineare Folge von Dipolmagneten, die in abwechselnder Nord-Süd-Ausrichtung hintereinander geschaltet sind (oft auch in so genannten *Halbach-Arrays*). Auf diese Weise werden die hindurchfliegenden Elektronen auf eine periodische, meist sinusförmige Bahn gezwungen, wobei sie Synchrotronstrahlung in ihrer durchschnittlichen Flugrichtung emittieren. Die dritte Generation von Synchrotronstrahlungsquellen ist durch die Entwicklung von *Undulatoren* gekennzeichnet. Bei dieser Weiterentwicklung des Wigglers wird durch die im Verhältnis kleinere Auslenkung der Elektronenbahn erreicht, dass es zur Interferenz aller Strahlungskegel kommt. Dies führt bei kleinerer relativer Photonenenergie zu einem schärferen Spektrum und erhöhter Brillanz. Die vierte Generation schließlich wird durch den so genannten *Elektronenlaser* (FEL) repräsentiert. Der FEL ist ein Undulator, der so (lang) gebaut ist, dass es zu einer Wechselwirkung zwischen der emittierten Strahlung und dem Elektronenpaket kommt. Der Abstand der Magnete und die Geschwindigkeit der Elektronen werden so aufeinander abgestimmt, dass die Lichtwellen, die an jedem einzelnen Magneten abgegeben werden, konstruktiv interferieren. Die Folge ist, dass die Intensität der Strahlung beim FEL nicht mehr linear, sondern proportional zum Quadrat der Anzahl emittierter Elektronen steigt.

### 3.4.3 Beugung von Röntgenstrahlen an Kristallen

Trifft Röntgenstrahlung auf Materie, kann es zur Wechselwirkung der Strahlung mit der (inneren) Elektronenhülle der Atome kommen, wobei diese ihrerseits Röntgenstrahlung in Form von kugelförmigen Wellen emittieren. Ändert sich dabei die ursprüngliche Wellenlänge, d.h. die Energie der Röntgenquanten nicht, so spricht man von der so genannten *Rayleigh-Streuung* bzw. der *klassischen Streuung*. Die von den einzelnen Atomen ausgehenden gebeugten Wellen interferieren miteinander, was eine konstruktive oder destruktive Interferenz, d.h. eine Verstärkung bzw. Abschwächung der gebeugten Strahlung zur Folge hat (Abbildung 3-7A). Da die Wellenlänge von Röntgenstrahlung im Bereich der Gitterparameter von Kristallen liegt, wirken diese als dreidimensionale Beugungsgitter. Die Vorausset-

zungen für konstruktive Interferenz bei der Beugung von Röntgenstrahlung am dreidimensionalen Gitter eines Kristalls lassen sich entweder durch die *Laue-Bedingungen* oder durch das äquivalente, graphisch anschaulichere *Bragg'sche Gesetz* beschreiben (Abbildung 3-7B).

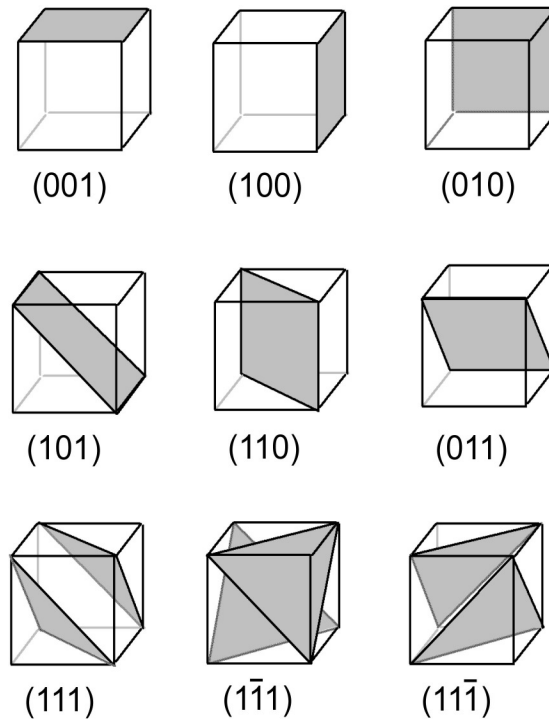


**Abbildung 3-7 Beugung von Röntgenstrahlen an Kristallen** (A) Interferenz elastisch gestreuter Röntgenstrahlung (B) Bragg'sches Gesetz

Bei der Bragg-Theorie werden die Atome bzw. Gitterpunkte des Kristalls gedanklich durch parallele Gitterebenen verbunden. Dabei entstehen Scharen so genannter *Netzebenen* mit konstantem Abstand. Diese Ebenen schneiden die zu den Kristallachsen parallel verlaufenden Gitterachsen  $a$ ,  $b$  und  $c$  bei  $x = a/h$ ,  $y = b/k$  und  $z = c/l$  (Abbildung 3-8). Die ganzzahligen Variablen  $h$ ,  $k$  und  $l$  werden als *Miller'sche Indizes* bezeichnet. Der Netzebenenabstand  $d_{hkl}$  und die „Belegung“ der Ebenen mit Gitterpunkten werden dabei mit steigenden Indizes immer kleiner. Wird monochromatische Röntgenstrahlung der Wellenlänge  $\lambda$  an einer Schar paralleler Netzebenen ( $hkl$ ) mit dem Abstand  $d_{hkl}$  gespiegelt (Einfallswinkel  $\theta$  = Ausfallswinkel), gilt nach dem Bragg-Gesetz folgender Zusammenhang:

$$2d_{hkl} \times \sin \theta_{hkl} = n \times \lambda \quad \text{Gleichung 3-1}$$

Für jede Netzebenenschar gibt es folglich für die verschiedenen Beugungsordnungen  $n$  einen Winkel  $\theta$ , den so genannten *Bragg- oder Glanzwinkel*, für den konstruktive Interferenz und damit ein *Reflex*  $hkl$  zu beobachten ist. Bei Kenntnis der Elementarzelle eines Kristalls ist man in der Lage, alle möglichen Netzebenen ( $hkl$ ) zu konstruieren und über ihre Netzebenenabstände  $d_{hkl}$  die Beugungswinkel der zugehörigen Reflexe  $hkl$  zu berechnen. Zwischen  $d_{hkl}$  und dem Glanzwinkel  $\theta$  besteht der reziproke Zusammenhang  $d_{hkl} \sim 1/(\sin \theta_{hkl})$ .



**Abbildung 3-8 Miller'sche Indizes im kubischen Kristallsystem** Der Koordinatenursprung liegt an der linken vorderen Ecke: x-Achse nach rechts, y-Achse in die Bildebene, z-Achse nach oben.

### 3.4.4 Reziprokes Gitter und Ewald-Konstruktion

Da die Darstellung der einzelnen Netzebenen im realen Raum schnell unübersichtlich wird, hat sich in der Kristallographie eine alternative Schreibweise etabliert. Dazu wird jede Netzebenenschar durch einen Vektor beschrieben, der die Richtung ihrer Flächennormale und die Länge des Netzebenenabstands besitzt. Die Endpunkte des Vektors sind durch die Achsenabschnitte  $a/h$ ,  $b/k$  und  $c/l$  gegeben. Führt man nun die reziproken Größen  $a^* = 1/a$ ,  $b^* = 1/b$ ,  $c^* = 1/c$  und  $d_{hkl}^* = 1/d_{hkl}$  ein, ergibt sich für den einfachen Fall eines orthogonalen Gittersystems der folgende Zusammenhang:

$$d_{hkl}^{*2} = h^2 a^{*2} + k^2 b^{*2} + l^2 c^{*2}$$

**Gleichung 3-2**

Die reziproken Einheiten  $a^*$ ,  $b^*$ ,  $c^*$  stellen demnach die Einheitsvektoren eines neuen Koordinatensystems dar, in dem sich der reziproke Netzebenenabstandsvektor  $d_{hkl}^*$ , auch *Streuvektor* genannt, durch die ganzzahligen Koordinaten  $h$ ,  $k$ ,  $l$  beschreiben lässt. Die Endpunkte der Vektoren  $d_{hkl}^*$  bilden schließlich das *reziproke Gitter*, wobei jeder Gitterpunkt einem möglichen Reflex  $hkl$  entspricht.

Die Einführung des reziproken Gitters ermöglicht die Veranschaulichung des Bragg'schen Gesetzes als Vektordiagramm. Diese Darstellung wird als *Ewald-Konstruktion* bezeichnet und ist besonders dafür geeignet, die praktische Durchführung von Beugungsexperimenten zu beschreiben. Zur Ableitung der Ewald-Konstruktion wird zunächst ein Kreis mit dem Radius  $1/\lambda$  und dem Mittelpunkt K konstruiert. Man bezeichnet diesen Kreis als *Ewald-Kreis* und seine dreidimensionale Projektion als *Ewald-Kugel* (Abbildung 3-9A). Die Verlängerung der auf den Kristall treffenden monochromatischen Röntgenstrahlung schneidet die Kugel im Ursprung O des reziproken Gitters. Für Strahlung, die an K unter dem Glanzwinkel  $2\theta$  zum Primärstrahl gebeugt wird, ist ein Reflex im Punkt P zu beobachten. Der Vektor  $\overrightarrow{KP}$  gibt dann die Richtung des gebeugten Strahls an,  $\overrightarrow{OP}$  entspricht dem Streuvektor  $d_{hkl}^*$ . Die Bragg'sche Gleichung lässt sich demnach in der folgenden Form schreiben:

$$\sin \theta = \frac{d_{hkl}^* / 2}{1 / \lambda} \quad \text{Gleichung 3-3}$$

Eine bestimmte Netzebene mit Abstand  $d_{hkl}$  verursacht also genau dann eine Reflexion, wenn der zugehörige Streuvektor  $d_{hkl}^*$  auf einer Kugeloberfläche mit dem Radius  $1/\lambda$  um K endet. Durch Drehung des Kristalls im Röntgenstrahl dreht sich auch das reziproke Gitter. Auf diese Weise schneiden nach und nach alle reziproken Gitterpunkte die Ewald-Kugel. Die maximale Anzahl der Reflexe, die durch Rotation des Kristalls beobachtet werden können, lässt sich mit Hilfe der Ewald-Kugel anschaulich ableiten (Abbildung 3-9B) und ergibt zu:

$$N = \frac{(4\pi/3) \times (2/\lambda)^3}{V_{\text{reziprok}}} = \frac{33.5 \times V}{\lambda^3} \quad \text{Gleichung 3-4}$$

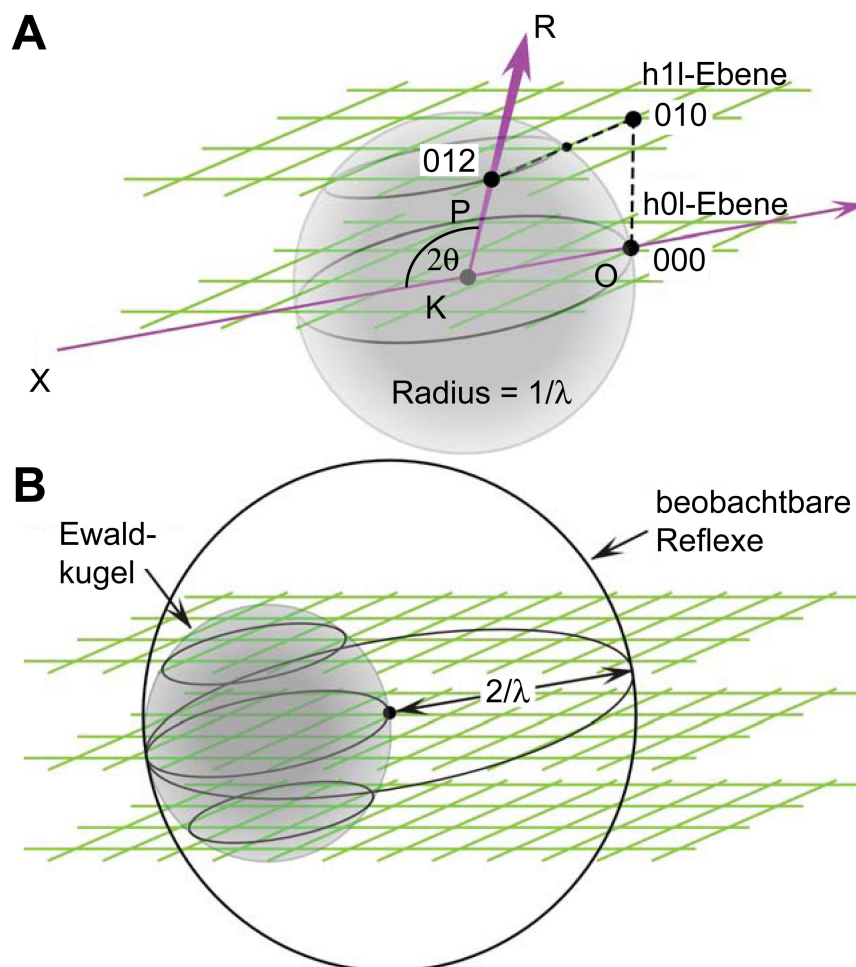
mit  $V_{(\text{reziprok})} \triangleq \text{Volumen der (reziproken) Einheitszelle}$

Das reziproke Gitter besitzt im Vergleich zum realen Kristallgitter ein zusätzliches Inversionszentrum (O). Im Beugungsbild kann daher für jeden Reflex  $hkl$  ein invertierter Reflex  $\bar{h}\bar{k}\bar{l}$  beobachtet werden, der unter dem Winkel  $-2\theta$  entstanden ist und die exakt gleiche Intensität aufweist. Es gilt also:

$$I_{hkl} = I_{\bar{h}\bar{k}\bar{l}} \quad \text{Gleichung 3-5}$$

Dieses Phänomen wird als *Friedel'sches Gesetz* bezeichnet und hat zur Folge, dass bei der Messung der Reflexintensitäten nur die Hälfte des reziproken Raums erfasst werden muss.

Durch die Kristallsymmetrie kann sich die Zahl der Reflexe, die zur vollständigen Vermessung eines Kristalls erfasst werden müssen, weiter verringern. Jene Reflexe, die sozusagen als „asymmetrische Einheit“ des reziproken Gitters die vollständige Beugungsinformation eines Kristalls tragen, nennt man auch die *unabhängigen* oder *einzigartigen Reflexe* (engl.: unique reflexions). Die übrigen Reflexe, die durch die Kristallsymmetrie erzeugt werden, werden als symmetrieäquivalente Reflexe bezeichnet. In der Praxis erhöht es die Genauigkeit einer Strukturbestimmung, wenn auch symmetrieäquivalente Reflexe gemessen werden. Die unabhängigen Reflexe werden dann durch Mittelung berechnet und den weiteren Strukturrechnungen zugrunde gelegt. Das Friedel'sche Gesetz gilt streng genommen nur für zentrosymmetrische Kristalle. Bei nicht-zentrosymmetrischen Kristallen treten geringfügige Unterschiede in den Intensitäten der *Friedelspaare*  $hkl$  und  $\bar{h}\bar{k}\bar{l}$  auf, die bei normaler Röntgenbeugung jedoch kaum ins Gewicht fallen und keinen Einfluss auf die Bestimmung der Lauegruppe haben. Das Friedel'sche Gesetz behält somit zumindest näherungsweise Gültigkeit.



**Abbildung 3-9 Die Ewald-Konstruktion** (A) Ewald-Kugel (B) Zur Ableitung der Anzahl der beobachtbaren Reflexe  $N$ . [Rhodes 2006]

### 3.4.5 Detektion gebeugter Röntgenstrahlung

In der Proteinkristallographie kommen zur Detektion der gebeugten Röntgenstrahlung meist Flächendetektoren wie *image plates* (IP) (dt.: Bildplatten) oder *CCD-Detektoren* (*charge coupled devices*) zum Einsatz.

#### Image Plate

Die IP besteht aus einer Trägerplatte oder -folie, die mit einer Beschichtung aus  $\text{Eu}^{2+}$ -dotiertem BaFCl überzogen ist. Durch auftreffende Röntgenquanten erfolgt photoinduzierte Oxidation von  $\text{Eu}^{2+}$  zu  $\text{Eu}^{3+}$  und ein Elektron wird ins Leitungsband abgegeben, welches an einer Gitterleerstelle, dem so genannten Farbzentrum, eingefangen wird (Halbwertszeit ca. 10 h). Die Anzahl der Farbzentren ist dabei proportional zur Intensität der einfallenden Strahlung. Nach der Belichtung wird die IP mit dem roten Licht eines Helium-Neon-Lasers bestrahlt, wodurch die metastabilen Farbzentren unter Emission von blauem Licht wieder in den Grundzustand zurückkehren. Die dabei emittierte Strahlung wird durch einen Photomultiplier detektiert. Vor jeder weiteren Aufnahme wird die IP durch Bestrahlung mit intensivem weißem Halogenlicht vollständig gelöscht. Der gesamte Auslesevorgang dauert in der Regel 1-2 min.

#### CCD-Detektoren

CCD-Detektoren werden seit Anfang der 1990er Jahre in der Proteinkristallographie eingesetzt [Naday *et al.* 1994]. Sie bestehen im Wesentlichen aus drei Komponenten: einem Phosphorbildschirm, einem Bildkonverter und einem Halbleiterzähler, dem eigentlichen CCD-Chip. Die einfallende Röntgenstrahlung wird zunächst an einem fluoreszierenden Phosphorbildschirm in sichtbares Licht umgewandelt. Über ein konisch zulaufendes Bündel von Glasfaserkabeln, den Bildkonverter, werden die Photonen an den CCD-Chip weitergeleitet. Das Licht trifft dort auf einen dotierten Silizium-Kristall, der proportional zur Photonenzahl elektrische Ladungen in Gestalt frei beweglicher Elektronen erzeugt, indem Valenzelektronen ins Leitungsband angeregt werden und so genannte Elektronen-Loch-Paare entstehen. Die in den einzelnen Sensorzellen entstandenen unterschiedlich großen Ladungspakete werden nach der Belichtung durch periodische Potentialänderungen horizontal durch den Siliziumkristall geschoben, von einem Auswerteverstärker in eine proportionale Spannung umgesetzt und diese schließlich detektiert. Die herausragende Eigenschaft von CCD-Detektoren ist neben ihrer hohen Empfindlichkeit die extrem schnelle

Detektoren ist neben ihrer hohen Empfindlichkeit die extrem schnelle Auslesezeit von nur wenigen Sekunden. Sie bieten zwar meist weniger aktive Messfläche als konventionelle Flächendetektoren und sind vergleichsweise teuer, kommen dafür aber ohne störanfällige Mechanik aus.

Im Rahmen dieser Arbeit wurde eine an eine Cu-Drehanode (AXS FR591, Bruker) gekoppelte MAR345-imaging plate (MAR Research) und CCD-Detektoren verschiedener Bauart an Synchrotron-Messplätzen des Deutschen Elektronen Synchrotron (BW6 und X13, DESY, Hamburg, Deutschland) und der European Synchrotron Radiation Facility (ID14-1 und ID23-2, ESRF, Grenoble, Frankreich) verwendet.

### 3.4.6 Strukturfaktorgleichung, Elektronendichte- und Patterson-Funktion

Wird Röntgenstrahlung an einer Netzebene  $hkl$  eines Kristalls gebeugt, so entspricht die resultierende Welle  $F_{hkl}$  der Summe aller Teilwellen, die an den Atomen  $j$  der Einheitszelle gebeugt wurde. Die Summe der einzelnen Streubeiträge kann durch die folgende *Strukturfaktorgleichung* beschrieben werden:

$$F_{hkl} = \sum_j^N f_j \exp[2\pi i(hx_j + ky_j + lz_j)] \exp[-B_j \sin^2 \theta / \lambda^2] \quad \text{Gleichung 3-6}$$

Die Größe  $f_j$  entspricht dabei der Amplitude des jeweiligen Beitrags und wird als *atomarer Streu- oder Formfaktor* bezeichnet. Er ergibt sich in erster Linie aus der Anzahl der (inneren) Elektronen  $Z$  des jeweiligen Atoms  $j$  und deren radialer Dichteverteilung. Er ist abhängig von der Wellenlänge  $\lambda$ , vom Bragg'schen Beugungswinkel  $\theta$  und damit vom Netzebenenabstand  $d_{hkl}$ . Dabei nimmt die Streuamplitude  $f_j$  mit zunehmendem Beugungswinkel  $\theta$  bzw. geringer werdendem Netzebenenabstand  $d_{hkl}$  ab (bei geringer Auflösung entspricht  $f_j$  der Anzahl der Elektronen). Die Werte für  $f_j$  sind in Abhängigkeit von  $\theta$  für fast alle Atome und Ionen in den *International Tables for Crystallography* tabellarisch erfasst und in den einschlägigen Programmen für die Röntgenstrukturanalyse bereits enthalten. Die fraktionellen Koordinaten  $x_j$ ,  $y_j$ ,  $z_j$  geben die Position des Atoms  $j$  in der Einheitszelle an und bestimmen somit die Phasenwinkel der resultierenden Wellen.  $h$ ,  $k$ ,  $l$  entsprechen den Frequenzen der Welle in Richtung  $x_j$ ,  $y_j$ ,  $z_j$  und geben die Indizes der jeweiligen Reflexe im reziproken Raum an.

Der zweite Term der Gleichung repräsentiert den so genannten *Auslenkungsparameter*. Er trägt der Tatsache Rechnung, dass die Atome eines Kristalls thermischen Schwingungen unterliegen und auf Grund der Unvollkommenheit des Kristallgitters von ihren Ideallagen abweichen. Die auch als *Debye-Waller-Faktor* oder *Temperaturfaktor* bezeichnete Größe  $B$  ist ein Maß für die daraus resultierende „Verschmierung“ der Elektronendichte der beugenden Atome. Wilson entwickelte 1949 eine Theorie, mit der sich ein mittlerer Temperaturfaktor aus den gemessenen Reflexintensitäten ableiten lässt. Hierzu werden zunächst in Abhängigkeit von der Auflösung mittlere Intensitäten  $\langle I \rangle$  und mittlere quadratische Streufaktoren  $\langle f^2 \rangle$  bestimmt. Wird anschließend  $\log \langle I \rangle / \langle f^2 \rangle$  gegen  $(\sin \theta / \lambda)^2$  bzw.  $1/d^2$  aufgetragen, ergibt sich der so genannte *Wilson Plot*, wobei die Steigung der Ausgleichsgeraden dem negativen mittleren Temperaturfaktor  $-B$  (Einheit:  $\text{\AA}^2$ ) entspricht.

Alternativ zu Gleichung 3-6 kann man die Strukturfaktoren  $F_{hkl}$  auch mit Hilfe der Elektronendichteverteilung innerhalb des Volumens  $V$  der Einheitszelle beschreiben.  $F_{hkl}$  lässt sich dann als Fourier-Transformierte der *Elektronendichte*  $\rho(x, y, z)$  schreiben:

$$F_{hkl} = \int_x \int_y \int_z \rho(x, y, z) \exp[2\pi i(hx + ky + lz)] dx dy dz \quad \text{Gleichung 3-7}$$

bzw.

$$F_{hkl} = \int_V \rho(x, y, z) \exp[2\pi i(hx + ky + lz)] dV \quad \text{Gleichung 3-8}$$

Daraus ergibt sich dann durch inverse Fourier-Transformation die folgende *Elektronendichtefunktion*:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} \exp[-2\pi i(hx + ky + lz)] \quad \text{Gleichung 3-9}$$

Die Elektronendichteverteilung kann also durch eine Fourier-Summe beschrieben werden, in der jeder Term einem Strukturfaktor entspricht. Da es sich bei den Strukturfaktoren  $F_{hkl}$  um Wellenfunktionen handelt, gilt:

$$F_{hkl} = |F_{hkl}| \exp i\alpha(h, k, l) \quad \text{Gleichung 3-10}$$

Dabei repräsentieren  $|F_{hkl}|$  und  $\alpha$  Amplitude bzw. Phase der Streuwelle. Es ergibt sich somit:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| \exp[-2\pi i(hx + ky + lz) + i\alpha(h, k, l)] \quad \text{Gleichung 3-11}$$



Die Berechnung der Elektronendichteverteilung innerhalb der Einheitszelle ist das Ziel der röntgenographischen Strukturaufklärung. Sie ist Grundlage des späteren Strukturmodells. Experimentell zugänglich sind jedoch nur die Strukturfaktoramplituden  $|F_{hkl}|$ . Das Produkt aus den komplexen Strukturfaktoren  $F_{hkl}$  und den konjugiert komplexen Strukturfaktoren  $F_{hkl}^*$ , also das Quadrat der Strukturfaktoramplitude  $|F_{hkl}|^2$ , ist proportional zu den gemessenen Reflexintensitäten  $I_{hkl}$ :

$$I_{hkl} \sim F_{hkl} \times F_{hkl}^* = |F_{hkl}|^2 \quad \text{Gleichung 3-12}$$

Die Bestimmung des Phasenwinkels  $\alpha$  für die einzelnen Reflexe  $F_{hkl}$  stellt somit die zentrale Herausforderung in der Kristallographie dar und wird auch als *Phasenproblem* bezeichnet. Für seine Lösung, die so genannte *Phasierung*, wurde eine Reihe von Methoden etabliert:

- **Isomorpher Ersatz**
- **Multiple Anomale Dispersion**
- **Molekularer Ersatz**
- **Direkte Methoden**
- **Mehrstrahlmethode**

Im Rahmen dieser Arbeit wurden die Methoden *Molekularer Ersatz* und *Multiple Anomale Dispersion* verwendet, die üblicherweise nach ihrer jeweiligen englischen Übersetzung als *molecular replacement*, kurz MR, und *multi-wavelength anomalous dispersion*, MAD bezeichnet werden.

### Patterson-Funktion

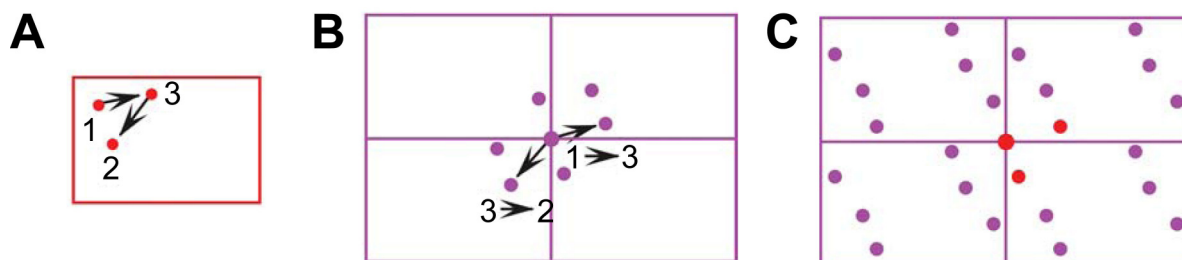
Die *Patterson-Funktion*  $P(u,v,w)$  ist eine Variation der Elektronendichtefunktion (**Gleichung 3-11**). Der Unterschied besteht in der Verwendung der Strukturfaktoramplituden-Quadrate als Fourier-Koeffizienten. Außerdem werden alle Phasenwinkel als null angenommen. Es ergibt sich somit:

$$P(u,v,w) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}|^2 \exp[-2\pi i(hu + kv + lw)] \quad \text{Gleichung 3-13}$$

$u,v,w$  sind die Gitterparameter der Patterson-Zelle und besitzen die gleiche Dimension wie  $x,y,z$  der „realen“ Elementarzelle. Die Maxima (Peaks) der Patterson-Funktion sind die Abstandsvektoren der  $N$  Atome in der Einheitszelle. Da jeder atomare Abstand in beiden Rich-

tungen gemessen auftaucht, ist die Patterson-Zelle zentrosymmetrisch und enthält insgesamt  $N^2$  Peaks.  $N$  Maxima, die so genannten *Eigenvektoren*, befinden sich im Ursprung bzw. Nullpunkt der Zelle, da jedes Atom zu sich selbst den Abstand null hat. Die relative Intensität eines Patterson-Maximums errechnet sich aus dem Produkt der Ordnungszahlen (Anzahl der Elektronen) der jeweiligen Atompaare.

Die Maxima der Patterson-Funktion werden in Form von *Elektronendichtekarten (Maps)* graphisch dargestellt. Ein einfaches Beispiel für die Konstruktion einer solchen Patterson-Map ist in Abbildung 3-10 gezeigt. Kristallstrukturen einfacher Moleküle können direkt mit Hilfe der Patterson-Funktion bestimmt werden. Im Falle von Proteinstrukturen wird die Patterson-Map auf Grund der Vielzahl ähnlicher Atome allerdings sehr kompliziert und führt nicht direkt zum Strukturmodell. Allerdings werden die folgenden Ausführungen zeigen, dass die Patterson-Methode ein mächtiges Werkzeug bei der Lösung des Phasenproblems ist.



**Abbildung 3-10 Schema zur Konstruktion einer Patterson-Map [Rhodes 2006]** (A) Einheitszelle mit drei Atomen ( $N = 3$ ). (B) Man erhält die Patterson-Map, indem alle interatomaren Vektoren im Ursprung platziert werden; das Ende eines jeden Vektors gibt die Position eines „Patterson Atoms“ (lokales Maximum/Peak der Patterson-Funktion) an. (C) Vollständige Patterson-Map mit allen  $N(N-1)$  Maxima für alle Einheitszellen; das Maximum im Ursprung entspricht den  $N$  Eigenvektoren einer jeden Zelle.

### 3.4.7 Phasenbestimmung durch Molekularen Ersatz

Auf Grund statistischer Untersuchungen an löslichen, globulären Proteinen wird davon ausgegangen, dass es nur eine begrenzte Anzahl einzigartiger Faltungstypen gibt [Brenner *et al.* 1997]. Die beiden bedeutendsten Datenbanken für die strukturelle Klassifizierung von Proteinen, CATH (Protein Structure Classification) und SCOP (Structural Classification of Proteins), unterscheiden derzeit unabhängig voneinander zwischen 1100 und 1200 einzigartige Faltungstypen [Cuff *et al.* 2009; Murzin *et al.* 1995]. Offensichtlich ist die Struktur von Protei-

nen deutlich stärker konserviert als ihre Aminosäuresequenz. So besitzen Proteine mit einer Sequenzidentität von mehr als 30% mit hoher Wahrscheinlichkeit oft bereits eine vergleichbare dreidimensionale Struktur [Sander & Schneider 1991].

Beim Verfahren des Molekularen Ersatzes macht man sich diese Tatsache zu Nutze, indem die Phaseninformation für ein Protein unbekannter Struktur (*Zielstruktur*) mit Hilfe der bereits bekannten Struktur eines ähnlichen Proteins (*Referenzstruktur* oder *Suchmodell*) gewonnen wird. Tatsächlich kann in einigen Fällen bereits eine Sequenzidentität von nur 25% zum Erfolg führen [Giorgetti *et al.* 2005].

Der Begriff Molekularer Ersatz geht auf den Titel einer 1972 veröffentlichten Monographie zurück, in der einer der Entwickler dieser Methode, Michael G. Rossmann, frühere Publikationen auf diesem Gebiet zusammenfasste [Rossmann 1972]. Das Verfahren wurde seither stark weiterentwickelt [Rossmann 1990; Turkenburg & Dodson 1996; Vagin & Teplyakov 2000] und ist derzeit die am häufigsten verwendete Phasierungsmethode [Evans & McCoy 2008].

Bei der MR-Methode handelt es sich im Grunde um ein 6-dimensionales Problem, das formal in eine *Rotations-* und eine *Translationssuche* unterteilt werden kann. Das Grundprinzip ist dabei der Vergleich der Patterson-Funktionen von Referenz- und Zielstruktur. Das Verfahren geht zurück auf Arbeiten von Rossmann und Blow, die 1962 eine Methode zum Nachweis nicht-kristallographischer Symmetrie durch Anwendung der so genannten *Rotationsfunktion* vorstellten [Rossmann 1962].

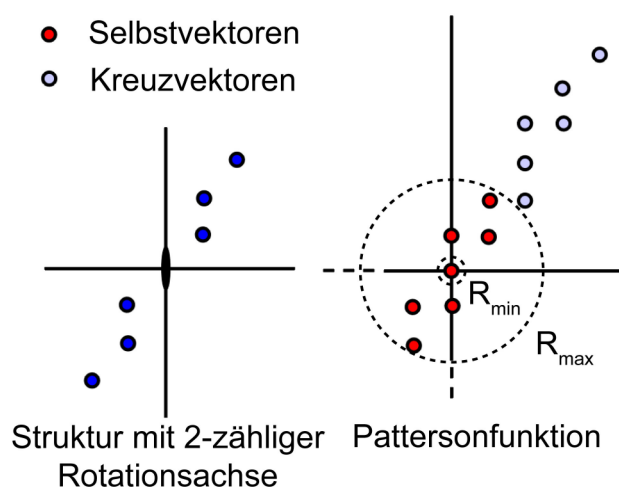


Abbildung 3-11 Selbst- und Kreuzvektoren in der Patterson-Map

## Rotationssuche

Die Patterson-Map ist eine Vektorkarte, in der die interatomaren Abstandsvektoren durch Peaks positiver Elektronendichte dargestellt werden. Für die Rotationssuche muss dabei zwischen zwei Klassen von Vektoren unterschieden werden: den *Selbstvektoren* zwischen Atomen desselben Moleküls und den *Kreuzvektoren* zwischen Atomen verschiedener Moleküle. Dies ist in Abbildung 3-11 für den Fall einer einfachen Struktur mit 2-zähliger Rotationssymmetrie gezeigt. In der Regel sind die meisten der Selbstvektoren kürzer als die Kreuzvektoren. Sie befinden sich in relativer Nähe zum Ursprung des Patterson-Raumes innerhalb eines Radius  $R_{\max}$ , welcher der Länge des größten Atomabstands innerhalb eines Moleküls entspricht. Das Prinzip des Molekularen Ersatzes basiert auf der Tatsache, dass sich die Länge der Selbstvektoren bei einer Rotation des Moleküls nicht ändert. In der Praxis wird die Patterson-Funktion für alle möglichen Orientierungen der Referenzstruktur R in einem Volumen U berechnet und mit der Zielstruktur Z verglichen. Dies geschieht mit Hilfe der Rotationsfunktion R:

$$R(M) = \int_U P_R(M \times u) \times P_Z(u) du \quad \text{Gleichung 3-14}$$

$P_R$  und  $P_Z$  sind dabei die Patterson-Funktionen der Referenz- bzw. Zielstruktur. Die Rotationsmatrix M beschreibt die Rotation des Ursprungsvektors u im Patterson-Raum U. Um sicherzustellen, dass die Referenz-Patterson-Funktion nur Selbstvektoren enthält, wird nur bis  $R_{\max}$  integriert. Bei langgestreckten Molekülen ist es außerdem sinnvoll, zusätzlich einen inneren Radius  $R_{\min}$  festzulegen, so dass es nicht zu einer Überlappung zwischen Referenz-Patterson und einem angrenzenden Ursprungs-Peak der Patterson-Funktion der Zielstruktur kommt. Das Volumen, über welches die Integration erfolgt, wird somit durch eine Hohlkugel um den Ursprung des Patterson-Raumes repräsentiert.

Die nachfolgende Gleichung stellt die 1962 von Rossmann und Blow vorgestellte Rotationsfunktion zur Auffindung nicht-kristallographischer Symmetrie in Form einer Fourier-Summentation dar (*Rossmann-Blow-Rotationsfunktion*):

$$R = \left( U/V^3 \right) \sum_z \sum_r |F_Z|^2 |F_R|^2 G_{z,z'} \quad \text{Gleichung 3-15}$$

Die Vektoren z und r geben dabei die Positionen von Ziel- bzw. Referenzstruktur an. V ist das Volumen der Einheitszelle. G stellt die so genannte Interferenzfunktion dar. Sie ist abhängig von z und z', wobei z' die Position des Vektors r nach Anwendung der Rotationsmatrix M angibt. Für ein sphärisches Integrationsvolumen mit Radius R ergibt sich  $G_{z,z'}$  zu:

$$G_{z,z'} = \frac{3(\sin 2\pi HR - 2\pi HR \times \cos 2\pi HR)}{(2\pi HR)^3} \quad \text{Gleichung 3-16}$$

mit  $H = |z + z'|$

Das im Rahmen dieser Arbeit benutzte MR-Programm MolRep [Vagin & Teplyakov 1997] verwendet eine Weiterentwicklung der ursprünglichen Rotationsfunktion [Rossmann 1972]. Bei dieser auch als *fast rotation function* (dt.: schnelle Rotationsfunktion) bezeichneten Variante werden die Patterson-Funktionen mit Hilfe von Kugelflächenfunktionen abgeschätzt. Die Rotationsfunktion wird dann mittels *schneller Fourier-Transformation* (engl.: *fast Fourier transform, FFT*) berechnet. Diese Vorgehensweise benötigt gegenüber der Berechnung der Rossmann-Blow-Rotationsfunktion bei gesteigerter Präzision deutlich weniger Rechenleistung (ca. Faktor 100). Außerdem vereinfacht sie die Berücksichtigung der Symmetrie der Patterson-Funktion und erleichtert die Entfernung der störenden Ursprungspeaks.

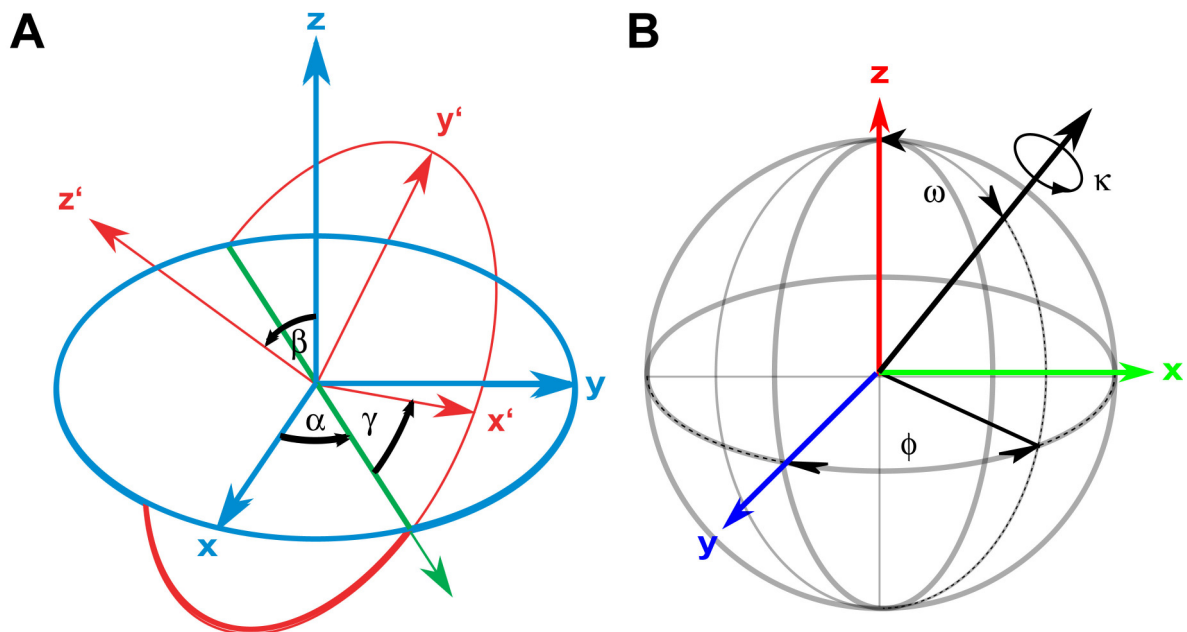
MolRep gibt die Ergebnisse der Rotationssuche in Form des Quotienten  $R/\sigma$  an, um die Maxima der Rotationsfunktion nach ihrer Signifikanz zu sortieren. Die Standardabweichung  $\sigma(R)$  berechnet sich dabei nach:

$$\sigma(R) = \left\{ \sum_n \frac{(R - \langle R \rangle)^2}{n} \right\}^{1/2} \quad \text{Gleichung 3-17}$$

mit  $\langle R \rangle = \text{Mittelwert der } n \text{ berechneten Rotationsfunktionen}$

### Winkeldefinitionen der Rotationsmatrix

Die Variablen der Rotationsmatrix werden üblicherweise durch zwei verschiedene winkelbasierte geometrische Systeme wiedergegeben (Abbildung 3-12). Am gebräuchlichsten ist die Verwendung der so genannten Euler'schen Winkel  $\alpha$ ,  $\beta$  und  $\gamma$ . Alternativ wird auch ein System mit den sphärischen Polarkoordinaten  $\omega$ ,  $\varphi$  und  $\chi$  genutzt. Letzteres System kommt vor allem bei der Berechnung so genannter *Eigenrotationsfunktionen* zum Einsatz. Bei diesem Verfahren zur Auffindung nicht-kristallographischer Symmetrie wird die Patterson-Funktion einer Struktur durch Rotation mit sich selbst überlagert. Da bei lokaler Symmetrie in Proteinkristallen vor allem zwei-, drei- oder vierzählige Drehachsen vorkommen, können für die Suche nach NCS die Werte des Polarwinkels  $\kappa$  auf  $180^\circ$ ,  $120^\circ$  oder  $90^\circ$  festgelegt werden.



**Abbildung 3-12 Winkeldefinitionen der Rotationsmatrix** (A) Beschreibung einer 3D-Rotation durch sukzessive Rotation um die Euler-Winkel  $\alpha$ ,  $\beta$  und  $\gamma$  (B) Kugelkoordinaten:  $\phi$  und  $\omega$  definieren eine Rotationsachse,  $\chi$  beschreibt die Rotation um diese Achse.

### Translationssuche

Wenn in der Rotationssuche eine mögliche Lösung gefunden wurde, gilt es die korrekte Position des Modells in der Einheitszelle zu bestimmen. Dies kann beispielsweise dadurch geschehen, dass für alle möglichen Positionen des vorläufigen Modells (und der symmetrieäquivalenten Moleküle) in der Einheitszelle so genannte *R-Faktoren*  $R$  oder *Korrelationskoeffizienten*  $KK$  bestimmt werden:

$$R = \frac{\sum_{hkl} \|F_{obs} - k F_{calc}\|}{\sum_{hkl} |F_{obs}|} \quad \text{Gleichung 3-18}$$

$$KK = \frac{\sum_{hkl} (|F_{obs}|^2 - \langle |F_{obs}|^2 \rangle) \times (|F_{calc}|^2 - \langle |F_{calc}|^2 \rangle)}{\left[ \sum_{hkl} (|F_{obs}|^2 - \langle |F_{obs}|^2 \rangle)^2 \sum_{hkl} (|F_{calc}|^2 - \langle |F_{calc}|^2 \rangle)^2 \right]^{1/2}} \quad \text{Gleichung 3-19}$$

mit:  $\langle |F|^2 \rangle$  bzw.  $\langle |F_{calc}|^2 \rangle$  = Durchschnittswerte der quadrierten experimentellen bzw. berechneten Strukturfaktoramplituden

Der Vorteil des Korrelationskoeffizienten gegenüber dem R-Faktor ist, dass seine Berechnung unabhängig vom Skalierungsfaktor  $k$  ist. Insgesamt benötigen derartige Trial & Error-Verfahren jedoch (auch heute noch) erhebliche Rechenkapazitäten, weshalb auch die Translationssuche für gewöhnlich durch Korrelation der Patterson-Funktionen für Ziel- und Referenzstruktur durchgeführt wird. Die *Translationsfunktion* hat die allgemeine Form:

$$T(t) = \int_U P_R(u, t) \times P_Z(u) du \quad \text{Gleichung 3-20}$$

$P_R$  und  $P_Z$  sind dabei wiederum die Patterson-Funktionen von Referenz- bzw. Zielstruktur. Die Vektoren  $t$  entsprechen den intermolekularen Vektoren zwischen den Molekülen der Einheitszelle. Ziel der Translationssuche ist die Ermittlung der relativen Positionen der Moleküle in der Einheitszelle. Diese Information liefern die Kreuzvektoren zwischen den Molekülen. Die Berechnung der Translationsfunktion wird daher möglichst nur unter Berücksichtigung der Kreuzvektoren von Ziel- und Referenzstruktur vorgenommen. Die von Crowther und Blow 1967 eingeführte so genannte T2-Translationsfunktion ist für  $n$  Moleküle in der Einheitszelle durch folgende Gleichung gegeben:

$$T_2(t) = \sum_h |F_Z(h)|^2 \sum_i^n \sum_j^n F_R(hM_i) F_R^*(hM_j) \exp(-2\pi i t_{ij}) \quad \text{Gleichung 3-21}$$

$|F_Z(h)|^2$  sind die gemessenen Intensitäten der Zielstruktur. Die Summation über  $i$  und  $j$  entspricht der Berechnung der Kreuzvektoren zwischen den symmetrieverwandten Molekülen  $i$  und  $j$  der Referenzstruktur  $R$ .  $M_i$  und  $M_j$  sind die jeweiligen Symmetrie-Operatoren und  $t_{ij}$  der intermolekulare Vektor zwischen den Molekülen. Um sicherzustellen, dass die T2-Funktion von Kreuzvektoren dominiert wird, können die Selbstvektoren der Referenzmoleküle von der Patterson-Funktion der Zielstruktur abgezogen werden.  $|F_Z(h)|^2$  in Gleichung 3-21 wird dann durch den folgenden Ausdruck ersetzt:

$$|F_Z(h)|^2 - k \sum_i^n |F_R(hM_i)|^2 \quad \text{Gleichung 3-22}$$

Der Faktor  $k$  ist dabei ein Koeffizient, der beobachtete und berechnete Intensitäten aufeinander skaliert. Die ursprüngliche T2-Funktion wurde im Laufe der Jahre beispielsweise durch die Normalisierung von Strukturfaktoren und Translationsfunktion sowie durch die Einführung zusätzlicher Terme wie der *Packungsfunktion* weiterentwickelt (z.B. [Harada *et al.* 1981]). Letztere bereinigt die Berechnung von falschen Maxima, die aus der physikalisch unsinnigen

Durchdringung der Moleküle der Referenzstruktur resultieren.

Auch das Programm MolRep verwendet eine Weiterentwicklung der  $T_2$ -Funktion, die mit einer Packungsfunktion  $Q$  korreliert ist.  $Q$  nimmt dabei Werte zwischen 0 und 1 ein, wobei  $Q = 1$  bedeutet, dass sich die Elektronendichten  $\rho$  der Moleküle  $i$  und  $j$  nicht durchdringen. Die zu den Symmetrie-Operatoren  $M_i$  und  $M_j$  gehörigen Terme der Translationsfunktion  $T_{ij}$  fließen dabei als Produkt und nicht als Summe in die Berechnung ein. Es ergibt sich somit:

$$T(t) = \prod_{i \neq k} T_{ij}(t) \times Q(t) \quad \text{Gleichung 3-23}$$

$$\text{mit } Q(t) = - \sum_{i \neq j} \int_U \rho_i(u, t) \rho_j(u, t) du \quad \text{Gleichung 3-24}$$

Zur Validierung der Ergebnisse der Translationssuche berechnet MolRep neben einem R-Faktor zwei weitere als *score* und *contrast* bezeichnete Größen. Der *score* entspricht dem Produkt aus Korrelationskoeffizient und dem Maximalwert der Packungsfunktion. Der Wert *contrast* wird aus dem Verhältnis des maximalen zum durchschnittlichen *score* gebildet. *Contrast*-Werte von  $> 1.3$  werden vom Programm als mögliche Lösungen akzeptiert. Werte  $> 2.5$  gelten als eindeutige Lösungen.

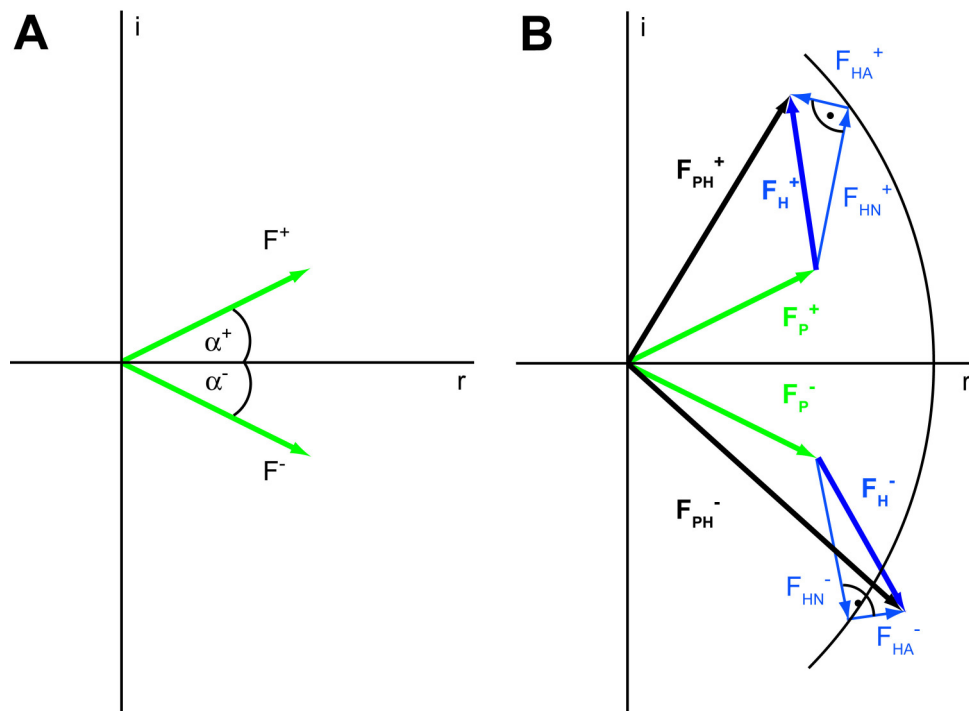
### 3.4.8 Phasenbestimmung durch multiple anomale Dispersion

Das in 3.4.4 eingeführte Friedel'sche Gesetz gilt nur für die in Proteinen am häufigsten vorkommenden Atome (C, N, O, S), deren Anregungsenergien erheblich oberhalb der Energie der gebräuchlichen Röntgenstrahlung liegen. Im Falle schwererer Atome, wie beispielsweise Se, kann eine der Absorptionskanten im Bereich der einfallenden Strahlung liegen. Ein Teil der auftreffenden Röntgenquanten führt dann zur Ionisation, was die ungerichtete Emission von z.B.  $K_\alpha$ -Strahlung des angeregten Elements zur Folge hat. Daraus resultiert eine erhöhte Untergrundstrahlung, und der restliche gestreute Anteil der Röntgenstrahlung erfährt infolge der stärkeren Wechselwirkung eine kleine Änderung in Amplitude und Phase. Dieses Phänomen wird als *anomale Streuung* oder *anomale Dispersion* bezeichnet. Um ihm Rechnung zu tragen, schreibt man den in 3.4.6 eingeführten atomaren Streufaktor  $f_i$  in allgemeiner Form:

$$f_i = f_0(\theta) + \Delta f'(\lambda) + i f''(\lambda) \quad \text{Gleichung 3-25}$$

Dabei stellt  $f_0$  den Anteil des Streufaktors weit weg von der Absorptionskante des anomalen Streuers dar. Er nimmt, wie wir bereits wissen, mit zunehmenden Glanzwinkel  $2\theta$  ab. Der





**Abbildung 3-13 Anomale Dispersion (A)** Strukturfaktoren eines Friedel-Paares ohne anomale Dispersion:  $F^+$  ist die Reflexion von  $F^-$  an der realen Achse des Argand-Diagrammes **(B)** Bruch des Friedel'schen Gesetzes auf Grund anomaler Dispersion in Gegenwart eines Schweratoms H.  $F_P^{+/-}$  und  $F_H^{+/-}$  entsprechen dem Protein- bzw. Schweratomanteil der Strukturfaktoren  $F_{PH}^{+/-}$ .  $F_{HN}^{+/-}$  und  $F_{HA}^{+/-}$  sind die normalen bzw. anomalen Beiträge zu  $F_H^{+/-}$ . Nähere Erläuterungen siehe Text.

Beitrag der anomalen Streuung ist von der Wellenlänge abhängig und setzt sich zusammen aus einem Realteil  $f'$  in Richtung  $f_0$  und dem dazu orthogonalen Imaginärteil  $f''$ , der stets positiv ist. Im Falle anomaler Streuung addiert sich also stets ein kleiner Phasenwinkel. Da es sich bei  $f'$  um den Wert handelt, um den der Anteil der normalen Streuung bei der Wellenlänge  $\lambda$  reduziert wird, lassen sich die Realteile des atomaren Streufaktors zusammenfassen, so dass vereinfacht geschrieben wird:

$$f_i = f' + if''$$

**Gleichung 3-26**

Für ein Protein P, das den anomalen Streuer H enthält, ergibt sich für die Strukturfaktoren eines Friedelpaares folgender in Abbildung 3-13 veranschaulichter Zusammenhang (die Beiträge der Reflexe  $hkl$  und  $\bar{h}\bar{k}\bar{l}$  sind durch + bzw. - gekennzeichnet):

$$F_{PH}^+ = F_P^+ + F_H^+ = F_P^+ + F_{HN}^+ + F_{HA}^+$$

$$F_{PH}^- = F_P^- + F_H^- = F_P^- + F_{HN}^- + F_{HA}^-$$

Gleichung 3-27

mit:  $F_{PH}^{+/-} \triangleq$  Strukturfaktoren des Proteins mit Schweratom  
 $F_P^{+/-} \triangleq$  Strukturfaktoren des Proteins  
 $F_H^{+/-} \triangleq$  Strukturfaktoren des Schweratoms  
 $F_{HN}^{+/-} \triangleq$  normaler Anteil der Strukturfaktoren des Schweratoms  
 $F_{HA}^{+/-} \triangleq$  anomaler Anteil der Strukturfaktoren des Schweratoms

Daraus ergibt sich die so genannte Bijvoet-Differenz zu:

$$F_{PH}^+ - F_{PH}^- = 2F_{HA}$$

Gleichung 3-28

$|F_{PH}^+|$  und  $|F_{PH}^-|$  sind experimentell zugängliche Größen.  $F_{HA}$  lässt sich daher mit Hilfe tabellierter Werte für  $f'$  und  $f''$  bei der entsprechenden Wellenlänge berechnen, wenn die Positionen der anomalen Streuer bekannt sind.

Für eine MAD-Phasierung werden in der Regel Datensätze bei drei unterschiedlichen Wellenlängen aufgenommen. Der für das Experiment wesentliche Bereich liegt dabei an der Absorptionskante des eingesetzten anomalen Streuers, da  $f'$  und  $f''$  dort ihre minimalen bzw. maximalen Werte erreichen. Der wellenlängenabhängige Verlauf von  $f'$  und  $f''$  ist in Abbildung 3-14 für das Element Selen dargestellt.

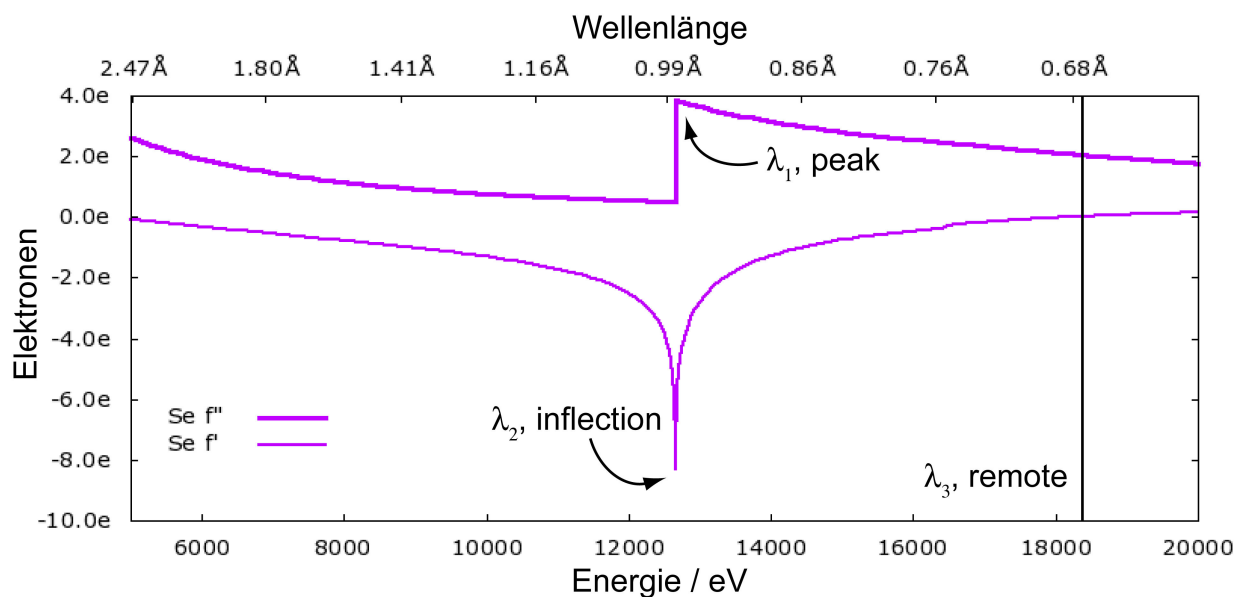


Abbildung 3-14 Wellenlängenabhängiger Verlauf von  $f'$  und  $f''$  für Selen

Zur optimalen Ausnutzung der Anteile von  $f'$  und  $f''$  wird am Maximum der  $f''$ -Kurve (*peak-Wellenlänge*,  $\lambda_1$ ) und am Minimum von  $f'$  bzw. am Wendepunkt von  $f''$  (*inflection-point-Wellenlänge*,  $\lambda_2$ ) gemessen. Darüber hinaus wird meist ein zusätzlicher Referenz-Datensatz weit weg von der Absorptionskante gemessen (*remote-Wellenlänge*,  $\lambda_3$ ). Die Positionen der anomalen Streuer können mit Hilfe der bereits erwähnten Patterson-Methode ermittelt werden. Hierzu werden zunächst die Differenzen der Bijvoet-Amplituden gebildet und mit dem Faktor  $f'/2f''$  skaliert. Diese anomalen Differenzen werden nun als Koeffizienten in die Patterson-Funktion (*anomale Differenz-Patterson-Funktion*) eingesetzt, welche dann die Positionen der anomalen Streuer liefert:

$$P(u, v, w) = \frac{1}{V} \sum_h \sum_k \sum_l \left\{ \left( |F_{PH}^+| - |F_{PH}^-| \right) \frac{f'}{2f''} \right\}^2 \exp[-2\pi i(hu + hv + hw)] \quad \text{Gleichung 3-29}$$

Mit Hilfe der Atompositionen können nun die Strukturfaktoren der Schweratome  $F_{H^{+/-}}$  berechnet werden. Da die Amplituden der Strukturfaktoren  $|F_{PH^{+/-}}|$  aus der Datensammlung bekannt sind und weiterhin der folgende Zusammenhang gilt (vgl. Gleichung 3-27):

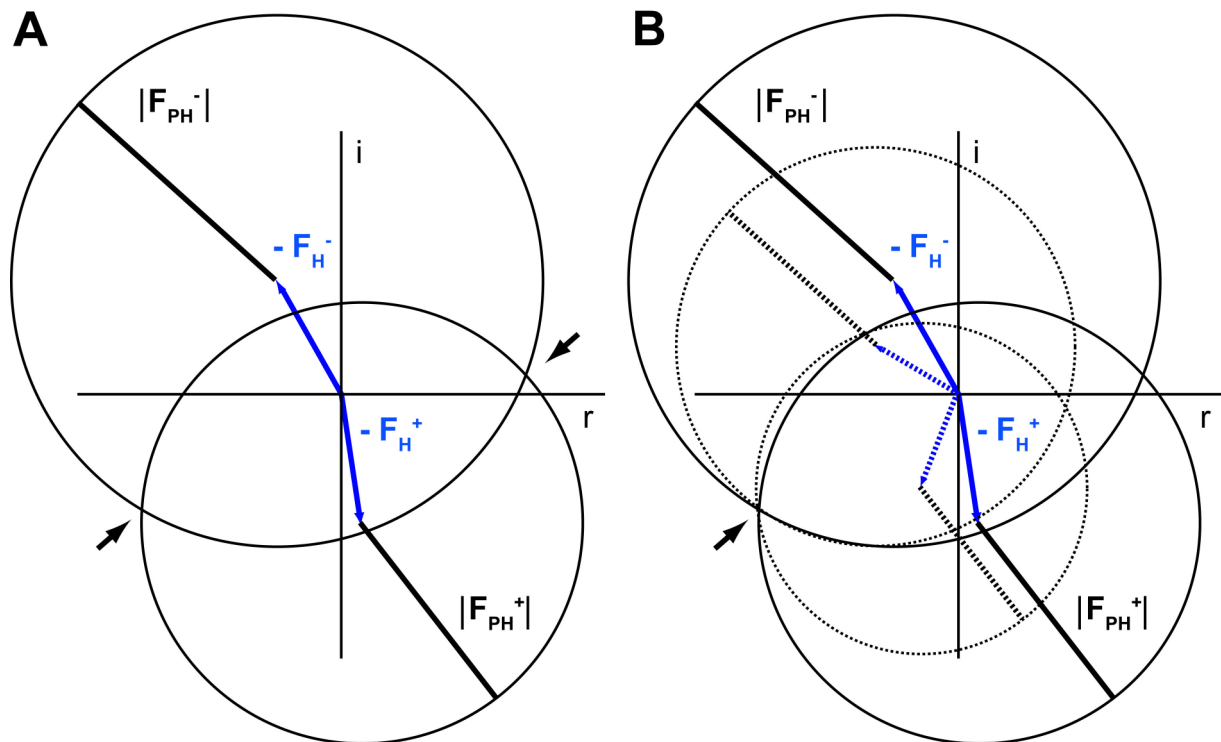
$$\begin{aligned} F_P^+ &= F_{PH}^+ - F_H^+ \\ F_P^- &= F_{PH}^- - F_H^- \end{aligned}$$

können die Phasen der Strukturfaktoren des nativen Proteins  $F_P^{+/-}$  mit Hilfe der *Harker-Konstruktion* bestimmt werden (Abbildung 3-15A) [Harker 1956]. Dabei ist zu beachten, dass sich erst bei Verwendung der Daten von zwei verschiedenen Wellenlängen (peak und inflection-point) eine eindeutige Lösung ergibt (Abbildung 3-15B).

Der Vorteil der MAD-Methode besteht darin, dass die gemessenen Datensätze in der Regel isomorph sind, da sie am gleichen Kristall gemessen wurden. Beinhaltet das Zielprotein keine anomal streuenden Atome, ist es besonders im Falle rekombinanter Produktion oft relativ problemlos möglich, anomal streuendes Selen in Form von Selenomethionin zu inkorporieren. Die Kenntnis der Selenomethionin-Positionen als „Ankerpunkte“ in der Elektronendichte kann überdies den anschließenden Bau des Modells entscheidend vereinfachen, da ihre Positionen in der Sequenz des Proteins bekannt sind.

Für die Lösung des Phasenproblems mit Hilfe der MAD-Methode wurden im Rahmen dieser Arbeit die Programme SOLVE und RESOLVE ([www.solve.lanl.gov](http://www.solve.lanl.gov)) [Terwilliger 2004] verwendet. Die Programme arbeiten nach dem parametrischem Schätzverfahren der *maximum-likelihood-Methode* (dt.: maximale Wahrscheinlichkeit) [Pratt 1976], das ausführlich in

[Terwilliger 1994] beschrieben ist. Nachfolgend soll lediglich die prinzipielle Vorgehensweise des Programms SOLVE dargestellt werden.



**Abbildung 3-15 Harker-Konstruktion zur Phasenbestimmung** Ausgehend vom Ursprung wird zunächst  $-F_H^+$  aufgetragen. Anschließend wird ein Kreis mit Radius  $|F_{PH}^+|$  und dem Mittelpunkt an der Spitze des Vektors  $-F_H^+$  eingezeichnet. Der Kreis repräsentiert alle Werte des Vektors  $F_{PH}^+ - F_H^+$  im Bezug auf alle möglichen Phasen von  $F_{PH}^+$ . Durch die analoge Konstruktion eines zweiten Kreises, erhält man alle möglichen Werte für  $F_{PH}^- - F_H^-$ . Die Schnittpunkte der Kreise zeigen die beiden Möglichkeiten für den Phasenwinkel an. (B) Mit Hilfe von MAD-Daten von zwei verschiedenen Wellenlängen ergibt sich eine eindeutige Lösung für den Phasenwinkel.

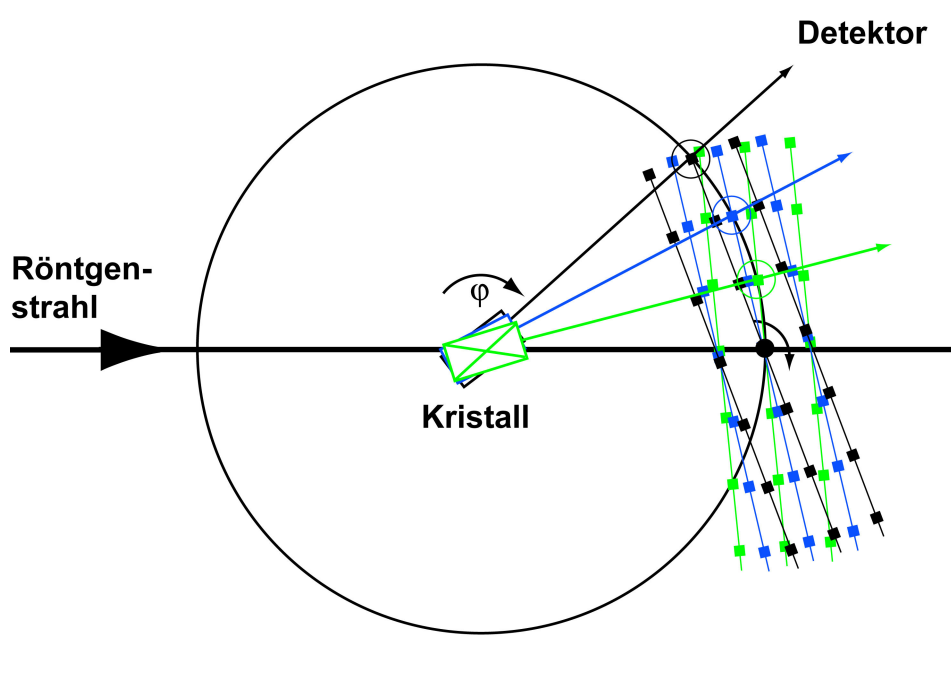
Der wichtigste Schritt einer MAD-Phasierung ist die Sammlung vollständiger, möglichst hoch aufgelöster Daten eines Selenomethionin-Derivatkristalls des Zielproteins, da der anomale Anteil der gemessenen Reflexintensität üblicherweise nur bei ca. 2-5% liegt [Walsh *et al.* 1999]. Daher muss auch die Prozessierung (Integration, Skalierung und Reduktion) dieser Daten mit besonderer Sorgfalt durchgeführt werden, da die zu ermittelnden anomalen Differenzen in Anbetracht des experimentellen Fehlers relativ gering sind. Der nächste Schritt besteht darin, die Positionen der Schweratome mit Hilfe der bereits erwähnten anomalen Differenz-Patterson-Funktion zu bestimmen. In SOLVE wird diese Suche automatisch von der Subroutine HASSP [Terwilliger *et al.* 1987] durchgeführt. Die gefundenen potenziellen Lösungen werden mit Hilfe eines speziellen Systems aus vier Einzelkomponenten bewertet:

- **Qualität der mit Hilfe der Schweratompositionen berechneten Elektronendichtekarte.** Das Kriterium ist hier das Auffinden zusammenhängender Regionen, die sich möglichst eindeutig als Solvens- bzw. Protein-Elektronendichte charakterisieren lassen.
- **Möglichst gute Übereinstimmung zwischen der experimentellen anomalen Differenz-Patterson-Funktion und der Funktion, die sich aus den Schweratompositionen errechnen lässt.**
- **Kreuzvalidierung mittels anomaler Fouriersynthese.** Dabei werden alle Schweratompositionen, bis auf eine, verwendet, um Phasen zu berechnen. Diese werden mit Hilfe der gemessenen anomalen Differenzen dazu benutzt, eine Elektronendichtekarte zu erzeugen, die für alle Schweratompositionen Maxima aufweisen sollte. Die Peakhöhe an der Position des weggelassenen Schweratoms ist ein Maß für die Glaubwürdigkeit dieser Position.
- **Figure-of-Merit (FOM).** Der FOM ist ein statistisches Maß für die Verlässlichkeit der Phasenberechnung. Er repräsentiert die Breite der Phasenwahrscheinlichkeitskurve  $P(\alpha)$  eines Reflexes. Der FOM entspricht ungefähr dem Kosinus des geschätzten Fehlers der berechneten Phase. Ein mittlerer FOM von 0.7 entspricht folglich einer Abweichung der gemittelten Phasen von ungefähr  $45^\circ$ .

Für jede der vier Komponenten wird ein so genannter *Z-Score* berechnet. Der höchste Gesamt-Z-Score zeigt dann die wahrscheinlichste Lösung an. Mit Hilfe der „Schwermetall(Sub)struktur“ kann anschließend eine erste Elektronendichte berechnet werden.

### 3.4.9 Datensammlung

Die röntgenographische Vermessung von Proteinkristallen wurde im Rahmen dieser Arbeit ausschließlich bei 100 K im Stickstoffstrom durchgeführt. Diese auch als *Kryo-Kristallographie* bezeichnete Vorgehensweise reduziert die strahlungsinduzierte Schädigung des Kristalls und den damit einhergehenden Qualitätsverlust der Messung erheblich. Der jeweilige Kristall wird mit Hilfe einer kleinen Nylonschleife (*Loop*) aus der Mutterlauge entnommen und in flüssigem Stickstoff schockgefroren. Unmittelbar vor dem Einfrieren wurden die Kristalle in der Regel in einen speziellen *Kryo-Puffer* überführt. Dieser entspricht möglichst genau der Zusammensetzung der Mutterlauge und beinhaltet zusätzlich Substanzen, so genannte *Kryo-Protetantien*, die die Bildung von störenden Eiskristallen verhindern sollen. Bewährt haben



**Abbildung 3-16 Messprinzip bei Flächendetektorsystemen.** Die Abbildung zeigt schematisch den auf dem Goniometerkopf befestigten Kristall von oben. Durch Drehung des Kristalls um das Winkelinkrement  $\Delta\varphi$  kommen nach und nach mehrere reziproke Gitterpunkte (Netzebenen) auf der Ewald-Kugel zum Liegen. Bei einer einzigen Aufnahme wird also eine Vielzahl von Reflexen registriert. Anschließend wird das Signal ausgelesen und der Kristall weitergedreht.

sich hierfür Glycerin, Ethylenglycol oder niedermolekulare Polyethylenglycole. In dieser Arbeit wurde Glycerin bei Konzentrationen von 15%-35% (v/v) verwendet. Der eingefrorene Kristall wurde anschließend auf einer genau justierbaren Halterung, dem so genannten *Goniometerkopf*, im Strahlengang der Röntgenquelle positioniert und während der Belichtung um einen kleinen Winkel  $\Delta\varphi$  (ca.  $0.3-1.0^\circ$ ) senkrecht zum einfallenden Röntgenstrahl gedreht (Abbildung 3-16).

Die entscheidenden Messparameter für eine erfolgreiche Datensammlung sind:

- Oszillationswinkel  $\Delta\varphi$
- Startwinkel  $\varphi_0$
- Winkelbereich der Messung  $\varphi_0 - \varphi_{\max}$  (Vollständigkeit der Messung)
- Belichtungszeit
- Kristall-Detektor-Abstand (maximale Auflösung)

Die optimalen Datensammelungsparameter wurden in der Regel mit dem Programmpaket MOSFLM [Leslie 2006] ermittelt. Die hierfür notwendige Bestimmung der Gitterparameter und der wahrscheinlichen Raumgruppe des Kristalls wird als *Indizierung* bezeichnet und

wurde mit der in MOSFLM implementierten *autoindexing*-Funktion durchgeführt. Dazu wurden jeweils mindestens zwei um  $90^\circ$  versetzte Aufnahmen ausgewertet. Die MOSFLM-Option *strategy* schätzt nach erfolgreicher Indizierung unter Berücksichtigung der Orientierung des Kristalls den für die maximale Vollständigkeit erforderlichen Winkelbereich ab.

Das Programm berechnet außerdem einen Wert für die *Mosaizität* des Kristalls. Diese auch als *Mosaikstruktur* bezeichnete Eigenschaft von Kristallen kann die Datensammlung erheblich beeinflussen. So entsprechen reale Kristalle nur in kleinen Bereichen dem idealen, durch dreidimensionales Aneinanderreihen von Elementarzellen beschriebenen Aufbau. Tatsächlich sind sie aus kleinen, zueinander leicht verkippten Idealkristallen in der Art eines Mosaiks aufgebaut. Die durchschnittliche Verkipfung dieser Bausteine wird in  $^\circ$  angegeben und liegt für Proteinkristalle von guter Qualität im Bereich  $0.2$  bis  $0.4^\circ$ . Auf Grund der Mosaikstruktur zeigt der reflektierte Strahl gegenüber dem einfallenden eine höhere Divergenz und reduzierte Kohärenz. Die Folge ist eine Verbreiterung der auf dem Detektor beobachteten Reflexe. Dies kann im schlimmsten Fall zu deren Überlappung führen und muss daher bei der Wahl des Oszillationswinkels berücksichtigt werden. Im Rahmen dieser Arbeit wurde in Winkelintervallen von  $\Delta\varphi = 0.3$ - $1.0^\circ$  pro Aufnahme gemessen. Die Mosaizität realer Kristalle ist allerdings auch eine notwendige Bedingung für die Gültigkeit von Gleichung 3-12, da der gebeugte Strahl den Kristall verlässt, ohne selbst Beugungseffekte und damit Intensitätsverluste durch Extinktion zu verursachen. Die Mosaikstruktur eines Kristalls kann in Einzelfällen durch mehrmaliges Auftauen und Einfrieren deutlich reduziert werden. Diese Technik wird als *crystal annealing* [Harp *et al.* 1998; Kriminski *et al.* 2002] bezeichnet.

MOSFLM wertet die einzelnen Aufnahmen einer Datensammlung aus und ordnet den Reflexen  $hkl$  eine gemittelte und um *Lorentz-* und *Polarisationsfaktor*<sup>1</sup> bereinigte Intensität  $I_{hkl}$  zu. Dieser Vorgang wird als *Integration* bezeichnet. Das Programm ordnet die Intensitäten in so genannte Auflösungsschalen an und gibt für diese jeweils den Quotienten aus den gemittelten Intensitäten  $I_{hkl}$  und deren Standardabweichung  $\sigma(I_{hkl})$  an. In der Regel wurde nach einer ersten Integration der Testaufnahmen der Kristall-Detektor-Abstand so gewählt, dass für die höchste Auflösungsschale ungefähr ein Wert von  $I_{hkl}/\sigma(I_{hkl}) = 2$  erreicht wurde. Weiterhin wurde vor Beginn der eigentlichen Vermessung des jeweiligen Kristalls die Belichtungszeit so eingestellt, dass es nicht zu einer Übersättigung der Reflexe insbesondere niedrigerer Auf-

---

<sup>1</sup> *Lorentz-Polarisationsfaktor*  $LP = (1 + \cos^2 2\theta)/(\sin^2 \theta \cos \theta)$

lösungschalen kam. Die Belichtungszeit pro Aufnahme betrug je nach Streukraft der Kristalle 10 bis 20 min für Cu-K $\alpha$ -Strahlung und lag bei der Verwendung von Synchrotronstrahlung im Bereich von wenigen Sekunden.

Eine zusätzliche Funktion von MOSFLM erlaubt es, Elementarzellparameter, Kristallorientierung, Messgeometrie und Mosaizität im Anschluss an die Integration im Rahmen eines automatischen *postrefinement* (dt.: „Nachverfeinerung“) mit Hilfe starker Reflexe zu verfeinern.

Für die Integration des hochaufgelösten nativen Enamidase-Datensatz wurde statt MOSFLM das weitgehend analoge Programm DENZO [Otwinowski & Minor 1997] verwendet.

Die Indizierung und die Festlegung der Messstrategie erfolgten am ESRF mit Hilfe des Programm-Pakets DNA [Leslie *et al.* 2002] größtenteils automatisch. DNA ist direkt mit der ESRF-Datenbank ISPYB (Information System for Protein CrystallographY Beamlines, <https://www.esrf.fr/ispyb/ispyb/security/logon.do>) verbunden und beinhaltet die folgenden Programme:

- MOSFLM: Integration der Daten
- BEST [Bourenkov & Popov 2006]: Berechnung der Messstrategie
- SCALA [Evans 2006]: Skalierung und Reduktion der Daten
- POINTLESS [Evans 2006]: Bestimmung der Raumgruppe

### 3.4.10 Skalierung und Reduktion der Daten

Die Intensitäten, die durch Integration der gesammelten Daten gewonnen wurden, unterliegen gewissen Ungenauigkeiten, die einer Reihe von experimentellen Schwankungen (schwankende Intensität des Röntgenstrahls, Strahlenschäden am Kristall, Anisotropie des Kristalls etc.) geschuldet sind. Es ist daher notwendig, die Intensitäten einer internen *Skalierung* zu unterziehen, so dass mehrfach gemessene Reflexe möglichst geringe Abweichungen vom Mittelwert ihrer Intensität aufweisen. Im Anschluss an die Skalierung werden die Intensitäten symmetrieäquivalenter Reflexe gemittelt und auf eine asymmetrische Einheit reduziert. Dieser Vorgang der *Datenreduktion* wird auch als *merging* bezeichnet.

Zur Skalierung und Datenreduktion der Datensätze wurde das Programm SCALA verwendet. Dieses gibt als Maß für die Datenqualität mehrere statistische Werte für die jeweiligen Auflösungsschalen an. Zur Veranschaulichung ist in Abbildung 3-17 ein Auszug aus der



Protokoll-Datei einer typischen Skalierung gezeigt.

Summary data for Project: gcda Crystal: gl2410_D3_01 Dataset: gl2410_D3_01			
	Overall	InnerShell	OuterShell
Low resolution limit	43.60	43.60	2.00
High resolution limit	1.90	6.01	1.90
Rmerge	0.054	0.027	0.279
Rmeas (within I+/I-)	0.061	0.031	0.318
Rmeas (all I+ & I-)	0.061	0.031	0.318
Rpim (within I+/I-)	0.029	0.015	0.151
Rpim (all I+ & I-)	0.029	0.015	0.151
Fractional partial bias	-0.030	-0.043	0.009
Total number of observations	215340	7088	30384
Total number unique	48670	1638	7059
Mean(I)/sd(I)	19.5	44.6	4.8
Completeness	100.0	99.6	100.0
Multiplicity	4.4	4.3	4.3

**Abbildung 3-17** Auszug aus der Protokoll-Datei der Skalierung eines GcdA-Datensatzes mit dem Programm SCALA [Evans 2006]

$R_{\text{merge}}$  ist dabei ein Maß für die Übereinstimmung aller, also auch mehrfach gemessener, symmetrieverwandter Reflexintensitäten:

$$R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i - \langle I \rangle|}{\sum_{hkl} \sum_i \langle I \rangle} \quad \text{Gleichung 3-30}$$

Dabei entspricht  $I_i$  der Intensität der  $i$ -ten Messung des Reflexes und  $\langle I \rangle$  dem Mittelwert aller für diesen Reflex gemessenen Intensitäten.  $R_{\text{merge}}$  wird häufig gleichbedeutend mit  $R_{\text{sym}}$  verwendet, allerdings handelt es sich bei Letzterem streng genommen um ein Maß für die Übereinstimmung symmetrieäquivalenter Reflexintensitäten ohne Mehrfachmessungen. Die *Multiplizität* (engl.: *multiplicity*) ist ein Maß für die Redundanz und damit die Qualität eines Datensatzes, sie entspricht dem Quotienten: (Gesamtzahl der gemessenen Reflexe)/(Anzahl der einzigartigen Reflexe). Die *Vollständigkeit* (engl.: *completeness*) eines Datensatzes definiert sich dementsprechend über das Verhältnis der Anzahl der gemessenen einzigartigen Reflexe zur Gesamtzahl der möglichen einzigartigen Reflexe.

$R_{\text{merge}}$  ist jedoch kein besonders zuverlässiges Maß für die Datenqualität, da er nur die Diskrepanz zwischen den beobachteten Reflexen angibt und nicht berücksichtigt, dass sich die Genauigkeit der gemittelten Intensitäten mit der Anzahl der einzelnen Messungen erhöht. In der Tat ist es so, dass  $R_{\text{merge}}$  mit steigender Multiplizität zunimmt. Seit einigen Jahren werden

daher alternative, von der Multiplizität unabhängige R-Faktoren wie  $R_{\text{meas}}$  und  $R_{\text{p.i.m.}}$  diskutiert [Diederichs & Karplus 1997; Weiss 2001; Weiss & Hilgenfeld 1997], die allerdings nach wie vor nicht sehr gebräuchlich sind:

$$R_{\text{meas}} = \frac{\sum_{hkl} \left( \frac{N}{N-1} \right)^{1/2} \sum_i |I_i - \langle I \rangle|}{\sum_{hkl} \sum_i |\langle I \rangle|} \quad \text{Gleichung 3-31}$$

$$R_{\text{p.i.m.}} = \frac{\sum_{hkl} \left( \frac{1}{N-1} \right)^{1/2} \sum_i |I_i - \langle I \rangle|}{\sum_{hkl} \sum_i |\langle I \rangle|} \quad \text{Gleichung 3-32}$$

mit  $N = \text{Multiplizität}$

$R_{\text{meas}}$  wird als multiplizitätsunabhängiger R-Faktor bezeichnet.  $R_{\text{p.i.m.}}$  ist ein Maß für die Präzision der gemittelten Messungen. Ein weiterer wichtiger Qualitätsindikator ist das Signal-zu-Rausch-Verhältnis  $\text{Mean}(I/\text{sd}(I))$ , es entspricht dem Mittelwert des Quotienten aus der mittleren Reflexintensität  $I$  und der zugehörigen Standardabweichung  $\text{sd}(I)$  (vgl. 3.4.9).

Der mit DENZO integrierte native Enamidase-Datensatz wurde mit dem Programm SCALEPACK [Otwinowski & Minor 1997] skaliert.

Im Anschluss an die Skalierung und Datenreduktion wurden die ermittelten Intensitäten jeweils mit dem Programm TRUNCATE [CCP4 1994] in Strukturfaktoramplituden umgerechnet.

Im Falle der SeMet-gelabelten Enamidase-Kristalle wurden für die MAD-Phasierung insgesamt drei Datensätze bei unterschiedlichen Wellenlängen gesammelt (vgl. 3.4.8). Zur möglichst genauen Bestimmung der anomalen Differenzen werden die Datensätze eines MAD-Experiments üblicherweise zunächst gemeinsam skaliert, wohingegen die anschließende Reduktion der Intensitäten, das *merging*, einzeln erfolgt. Scala gibt zur Abschätzung der Differenzen zwischen Amplituden der Bijvoet-Paare einen speziellen R-Faktor,  $R_{\text{anom}}$ , an:

$$R_{\text{anom}} = \frac{\sum_{hkl} |\langle I^+ \rangle - \langle I^- \rangle|}{\sum_{hkl} |\langle I^+ \rangle + \langle I^- \rangle|} \quad \text{Gleichung 3-33}$$

mit  $\langle I^+ \rangle / \langle I^- \rangle = \text{gemittelte Intensitäten der Bijvoet-Paare}$

### 3.4.11 Dichtemodifikation

Es wurde bereits darauf hingewiesen (vgl. 3.4.8), dass die Qualität der aus Schweratompositionen berechneten, experimentellen Phasen und der resultierenden Elektronendichtekarten oft nicht ausreicht, um den Verlauf der Polypeptidkette eindeutig zu erkennen und ein erstes Strukturmodell zu erstellen. Dies ist im Falle einer MAD-Phasierung zum Teil auf experimentelle Fehler bei der Bestimmung der geringen anomalen Differenzen zurückzuführen. Darüber hinaus weisen die zur Lösung der Substruktur verwendeten Daten des Schwermetall-Derivats meist eine deutlich geringere Auflösung auf als die Daten des nativen Proteins. In die Berechnung der ersten Elektronendichte fließt daher nur ein Teil der theoretisch verfügbaren Informationen ein. Vor dem eigentlichen Modellbau wird aus diesem Grund versucht, die experimentellen Phasen zu verbessern. Dieser Prozess wird als *Dichtemodifikation* bezeichnet. Durch Ausnutzung relevanter Eigenschaften der Zielstruktur sowie allgemeiner chemisch-physikalischer Prinzipien kann dabei die Elektronendichte in einem iterativen Prozess oft signifikant verbessert werden, ohne dass ein Strukturmodell notwendig ist. Nachfolgend werden die wichtigsten Methoden der Elektronendichtemodifikation skizziert.

#### Solvent flattening

Die erste automatisierte Methode des *solvent flattening* geht zurück auf Wang [Wang 1985] und wurde seither stetig weiterentwickelt [Abrahams 1997; Leslie 1987; Terwilliger 2000]. Hier soll nur das zu Grunde liegende Prinzip kurz erläutert werden.

Die Elementarzelle von Proteinkristallen enthält typischerweise große Bereiche, die mit ungeordneten Lösungsmittelmolekülen angefüllt sind (ca. 30-70%). Der Beitrag dieser Bereiche zur Elektronendichte ist vergleichsweise gering und annähernd konstant. Innerhalb der Proteinregion variiert die Elektronendichte wesentlich stärker. Daher können Lösungsmittel- und Proteinbereich über die jeweilige Standardabweichung  $\sigma$  der Elektronendichte innerhalb eines definierten Volumens (Gleichung 3-34) voneinander unterschieden werden.

$$\sigma = \sqrt{\frac{\sum (\rho - \langle \rho \rangle)^2}{n}} \quad \text{Gleichung 3-34}$$

mit  $\sigma$  = Standardabweichung der Elektronendichte innerhalb des betrachteten Volumens  
 $\rho$  = Elektronendichte am Gitterpunkt  
 $\langle \rho \rangle$  = mittlere Elektronendichte innerhalb des betrachteten Volumens  
 $n$  = Anzahl der Gitterpunkte innerhalb des betrachteten Volumens

Der Lösungsmittelregion kann auf diese Weise ein konstanter, niedriger Elektronendichtewert zugewiesen werden. Aus der modifizierten Dichte werden anschließend neue Phasen berechnet.

### **Histogram Matching [Lunin 1988]**

Systematische Untersuchungen an makromolekularen Röntgenstrukturen haben gezeigt, dass die Häufigkeitsverteilung einzelner Werte innerhalb der Elektronendichte bei einer bestimmten Auflösung und vergleichbarem Lösungsmittelgehalt weitgehend unabhängig vom betrachteten Protein ist. Eine Anpassung experimenteller Elektronendichten an diese theoretische Verteilung kann zu einer signifikanten Verbesserung schlechter Elektronendichtekarten führen. Dazu wird die Elektronendichte für alle Punkte eines Gitters berechnet, nach ihrer Höhe sortiert und die einzelnen Werte mit ihnen äquivalenten Werten aus der Standardverteilung ersetzt. Aus dieser verbesserten Elektronendichte können im Anschluss wiederum neue Phasen berechnet und mit den ursprünglichen Phasen kombiniert werden.

### **NCS averaging**

Diese Methode wurde zum ersten Mal 1974 von Bricogne erfolgreich eingesetzt [Bricogne 1974]. Das *NCS averaging* (dt.: NCS-Mittelung) beruht auf der Überlegung, dass die Elektronendichte der Kopien des Proteins, die durch NCS verknüpft sind, prinzipiell identisch sein sollte. Sind die entsprechenden Rotations- und im Falle offener NCS auch die Translationsooperatoren bekannt und steht ein vorläufiges Modell des Monomers (Maske) zur Verfügung, ist es möglich, die Elektronendichte über die symmetrieäquivalenten Bereiche zu mitteln und damit die experimentell bestimmten Phasen entscheidend zu verbessern. Die NCS-Operatoren lassen sich im Falle einer Phasierung durch MAD oder MR mit Hilfe der Schweratom-Koordinaten bzw. der Ergebnisse von Rotations- und Translationssuche ermitteln. Die Maske kann aus einem bereits zur Verfügung stehenden vorläufigen Modell (z.B. im Rahmen einer MR-Phasierung) berechnet werden. Nach erfolgreicher Phasierung durch Schweratom-Methoden kann sie aus der Substruktur oder aber nach *solvent flattening* direkt aus der Elektronendichte abgeleitet werden.

**wARP [Perrakis *et al.* 1997]**

Das Programm wARP (engl.: *weighted ARP*) ist eine Erweiterung des 1993 von Lamzin & Wilson vorgestellten ARP-Algorithmus (engl.: *automated refinement procedure*) [Lamzin & Wilson 1993] für die automatisierte Verfeinerung von Proteinstrukturmodellen. Die wARP-Methode ermöglicht es, kristallographische Phasen und die daraus resultierenden Elektronendichtekarten nach bereits erfolgter Dichtemodifikation durch z.B. solvent flattening oder NCS-averaging weiter zu verbessern. Dabei wird in die Ausgangselektronendichte zunächst eine große Anzahl von „Dummy-Atomen“ (z.B. Sauerstoffatome) eingebaut. Im Anschluss daran werden die Positionen durch einen automatisierten Prozess verfeinert und somit das Modell sukzessive verbessert. Dabei werden die Strukturverfeinerungsprogramme (siehe auch 3.4.12) PROLSQ [Hendrickson & Konnert 1980], REFMAC [Murshudov *et al.* 1997] und ARP in einem iterativ miteinander kombiniert. Ausgehend von diesem ersten Modell werden sechs weitere Modelle generiert. Die Mittelung über diese führt zu einer Reduzierung des Phasenfehlers und resultiert oft in einer signifikanten Verbesserung der Elektronendichte. Eine wichtige Voraussetzung für das Gelingen der wARP-Methode ist die Verfügbarkeit nativer Daten guter Auflösung (min. 2.4 Å).

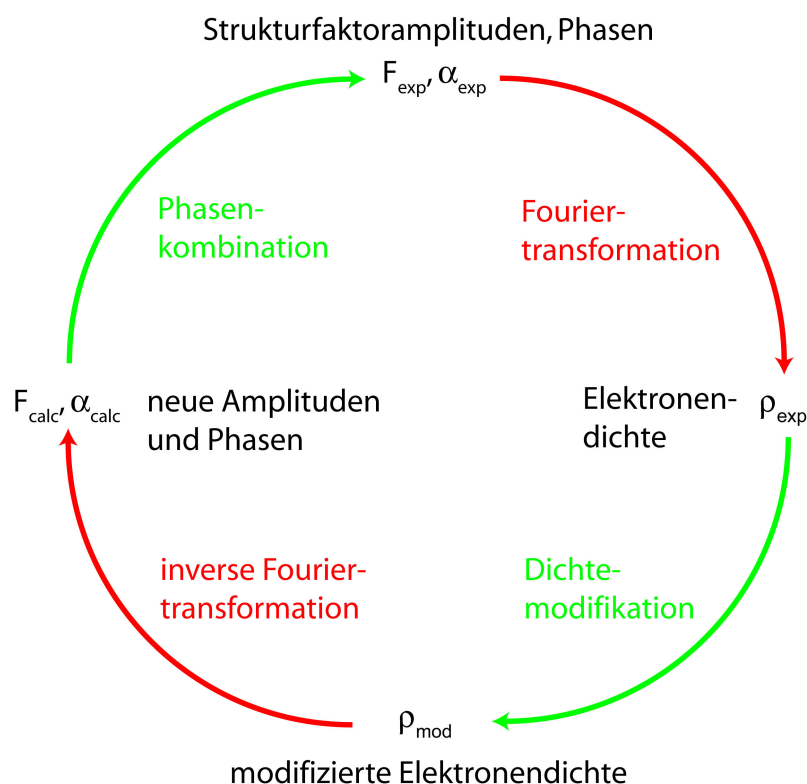


Abbildung 3-18 Prinzip der zyklischen Dichtemodifikation

Die hier beschriebenen Methoden der Dichtemodifikation werden in der einschlägigen Software häufig in iterativen Prozessen miteinander verknüpft. Das Grundprinzip der zyklischen Verfeinerung ist dabei in den meisten Fällen ähnlich (vgl. Abbildung 3-18). Die im Rahmen dieser Arbeit verwendeten Programme ARP/wARP [Lamzin & Wilson 1997] und RESOLVE [Terwilliger 2004] bieten überdies die Möglichkeit, experimentell bestimmte Phasen mit Hilfe nativer Daten höherer Auflösung über den ursprünglichen Auflösungsbereich hinaus zu erweitern.

### 3.4.12 Modellbau

#### Automatisierte Verfahren

Auch wenn Strukturamplituden und Phasen von ausreichender Qualität vorliegen, ist der manuelle Einbau des atomaren Modells in die Elektronendichtekarte ein zeitaufwendiger und mitunter sehr mühsamer Prozess. Daher wurden Methoden entwickelt, um diesen Vorgang mit Hilfe leistungsfähiger Rechner möglichst vollständig zu automatisieren. Die zugrunde liegenden Prinzipien basieren in vielen Fällen auf den vorgestellten Methoden der Dichtemodifikation. So beinhalten beispielsweise die Programmpakete SOLVE/RESOLVE und ARP/wARP bereits spezielle Algorithmen für den automatisierten Strukturmodellbau.

Im Rahmen dieser Arbeit konnte auf Grundlage der mit SOLVE/RESOLVE verfeinerten experimentellen Phasen und der Proteinsequenz ein nahezu vollständiges Strukturmodell des Enzyms Enamidase automatisch mit der in ARP/wARP implementierten *warpNtrace*-Funktion [Perrakis *et al.* 1999] erstellt werden. Die Methode ist im Wesentlichen eine Weiterentwicklung des wARP-Algorithmus (vgl. 3.4.11). Die „Dummy-Atomstruktur“ wird zunächst gemäß den geometrischen Einschränkungen, die für Proteine gelten, analysiert. Die resultierenden Peptid-Teilstränge werden dann mit dem „Dummy-Modell“ kombiniert, so dass ein Hybridmodell entsteht, welches auf die bereits in 3.4.11 skizzierte Weise iterativ verfeinert wird und in der nächsten „Runde“ als Ausgangspunkt für ein verbessertes Modell dient. Auf diese Weise werden die „Dummy-Atome“ nach und nach durch Peptide ersetzt. Dieser Teil des automatisierten Modellbaus beschränkt sich zunächst auf die Proteinhauptkette und wird daher als *mainchain tracing* (engl.: *to trace* = aufspüren) bezeichnet. Im Anschluss daran wird mit Hilfe der Proteinsequenz versucht, die Seitenketten einzubauen. Da ARP/wARP die Verfeinerungsroutine *ARP/waters* enthält, wird im Rahmen des Modellbauprozesses bereits ein großer Teil der an das Protein gebundenen Wassermoleküle automa-

tisch in die Elektronendichte eingebaut. Nach einer Reihe von Zyklen ist das Resultat im günstigsten Fall ein nahezu vollständiges atomares Strukturmodell des Zielproteins.

### **Molekularer Ersatz**

Wurde die Struktur mit Hilfe der MR-Methode gelöst, liegt oft schon ein relativ gutes Strukturmodell vor. Programme wie MolRep bieten die Möglichkeit, neben dem Referenzmodell auch die Proteinsequenz des Zielproteins einzulesen, so dass die Aminosäurereste der Referenzstruktur automatisch ausgetauscht werden können.

### **Isomorphe Kristallformen**

Handelt es sich bei der Zielstruktur um eine Kristallform, die zu einer bereits bekannten (nativen) Struktur isomorph ist, vereinfacht sich das Problem erheblich. Führt die Inkorporation des Liganden bei Kokristallisations- bzw. *soaking*-Experimenten zu keinen größeren Änderungen des Kristallgitters, ist es in der Regel möglich, die gemessenen Strukturfaktoramplituden direkt mit Hilfe der Koordinaten der nativen Struktur zu verfeinern.

### **Manueller Modellbau**

In der Realität liefern automatisierte Modellbauverfahren meist nur ein unvollständiges Bild der Proteinstruktur. Häufig enthält das vorläufige Modell noch mehr oder weniger ausgehende Fehlstellen, oder die Seitenketten einzelner Aminosäuren befinden sich noch nicht in Übereinstimmung mit der Elektronendichte. Bei niedriger Auflösung oder schwacher Datenqualität versagen automatische Verfahren mitunter vollständig. In solchen Fällen muss der Modellbau ganz oder teilweise manuell fortgesetzt werden.

Proteinstrukturen enthalten neben der/den Polypeptidkette/n noch weitere Moleküle. So befindet sich in direkter Nähe zum Protein in der Regel zusätzliche Elektronendichte, die geordneten Solvensmolekülen zugeordnet werden kann. Größtenteils handelt es sich dabei um Wasser, aber auch andere niedermolekulare Bestandteile der Kristallisationslösung, wie beispielsweise Glycerin-Moleküle, sind häufiger zu finden. Das Protein kann überdies Kofaktoren oder im Falle von Ligand-Protein-Komplexen Substrate oder Substrat-Analoga enthalten.

In dieser Arbeit wurde für den manuellen Modellbau das interaktive Grafikprogramm *Coot* [Emsley & Cowtan 2004] verwendet. Ähnlich wie das bereits erwähnte Programm ARP/wARP enthält Coot eine Funktion, die geordnete Wassermoleküle teilautomatisch in

die Dichte einbaut. Für die Verfeinerung des Strukturmodells (vgl. 3.4.13) ist es notwendig, dass für alle Moleküle eine so genannte *Library*-Datei zur Verfügung steht, die eine vollständige chemische und geometrische Beschreibung des jeweiligen Moleküls enthält. Für viele der in der Kristallisation von Biomolekülen relevanten Moleküle sind diese in Coot bereits vorhanden. Falls nicht, können sie entsprechenden Datenbanken (vgl. 3.4.14) entnommen oder aber mit Hilfe spezieller Programme manuell erstellt werden. Im Rahmen der vorliegenden Arbeit wurde dafür das Programm *Libcheck* in Verbindung mit der dazugehörigen graphischen Eingabeoberfläche *Sketcher* (beide [CCP4 1994]) verwendet. Die Ergänzung und Korrektur der Modelle wurden in Coot anhand folgender Differenz-Fourier-Elektronendichtekarten durchgeführt.

#### *(F<sub>obs</sub>-F<sub>calc</sub>)-Elektronendichtekarte*

*(F<sub>obs</sub>-F<sub>calc</sub>)-Differenzdichtekarten* werden nach der folgenden Gleichung berechnet:

$$\rho_{obs} - \rho_{calc} = \frac{1}{V} \sum_{hkl} (|F_{obs}| - |F_{calc}|) \exp[i\alpha_{calc}] \exp[-2\pi(hx + hy + hz)] \quad \text{Gleichung 3-35}$$

Für die Berechnung von  $\rho_{obs}$  werden also als Koeffizienten die gemessenen Strukturfaktoramplituden  $|F_{obs}|$  und die aus dem Modell berechneten Phasen  $\alpha_{calc}$  verwendet. Die Differenzdichte erscheint mit halber tatsächlicher Höhe und ist signifikant ab einem Differenzdichteniveau von etwa drei Standardabweichungen  $\sigma$  der gemittelten Elektronendichte, d.h. ab einem *Konturierungsgrad* von ca.  $3\sigma$ . Dabei zeigt positive Differenzdichte fehlende Atome an, während negative Differenzdichte auf falsch positionierte Bereiche hindeutet. *(F<sub>obs</sub>-F<sub>calc</sub>)-Karten* werden daher sowohl für den Einbau fehlender Aminosäuren, Wassermoleküle oder Liganden als auch zur Korrektur von Modellbereichen verwendet.

#### *(2F<sub>obs</sub>-F<sub>calc</sub>)-Elektronendichtekarte*

Die Berechnung der Dichte  $\rho_{obs}$  ist aufgrund der Verwendung der berechneten Phasen  $\alpha_{calc}$  relativ stark durch das Strukturmodell beeinflusst (engl.: *model bias*). Daher wird zusätzlich eine korrigierte Dichtekarte mit den Koeffizienten  $(2F_{obs}-F_{calc}) \exp(i\alpha_{calc})$  berechnet. Dies entspricht der Addition der  $|F_{obs}| \exp \alpha_{calc}$ -Karte mit der  $(|F_{obs}| - |F_{obs}|) \exp \alpha_{calc}$ -Differenzdichtekarte. Sie wird bei einem Konturierungsgrad von  $> 1\sigma$  dargestellt und zeigt neben der Elektronendichte des Modells die Differenzdichte in voller Höhe an.



*(mF<sub>obs</sub>-DF<sub>calc</sub>)- und (2mF<sub>obs</sub>-DF<sub>calc</sub>)-Elektronendichtekarten*

Unvollständige Strukturmodelle weisen meist noch erhebliche Koordinatenfehler auf. Elektronendichten, die auf Grundlage der aus dem Modell berechneten Phasen erstellt wurden, zeigen daher teils enormen *model bias*. Heutzutage werden daher häufig so genannte  $\sigma_A$ -gewichtete *(mF<sub>obs</sub>-DF<sub>calc</sub>)- und (2mF<sub>obs</sub>-DF<sub>calc</sub>)-Dichtekarten* verwendet [Read 1986]. Der Faktor m entspricht dabei dem FOM der Modellphasen, D ist ein von der Auflösung abhängiger Koeffizient, der Fehler in den atomaren Koordinaten des momentanen Modells widerspiegelt. Er wird aus der graphischen Auftragung des so genannten  $\sigma_A$ -Faktors gegen  $(\sin \theta/\lambda)^2$  ermittelt. Es gilt der folgende Zusammenhang:

$$\sigma_A = D \left( \frac{\sum_P P}{\sum_N N} \right)^{1/2} \quad \text{Gleichung 3-36}$$

Der Term  $(\sum_P/\sum_N)^{1/2}$  gibt dabei an, inwieweit das Strukturmodell P mit der Struktur N übereinstimmt. In dieser Arbeit kamen für Modellbau und -verfeinerung (3.4.13) fast ausschließlich *(mF<sub>obs</sub>-DF<sub>calc</sub>)- und (2mF<sub>obs</sub>-DF<sub>calc</sub>)-Elektronendichtekarten* zum Einsatz.

### 3.4.13 Verfeinerung des Strukturmodells

Ziel der Verfeinerung des Strukturmodells ist es, durch Variation der Positions- und der Auslenkungsparameter aller (Nicht-Wasserstoff-)Atome eine möglichst gute Übereinstimmung zwischen den gemessenen und den aus dem Modell berechneten Strukturfaktoren  $F_{\text{obs}}$  bzw.  $F_{\text{calc}}$  zu erreichen. Ein Maß für diese Übereinstimmung ist der *kristallographische R-Faktor*  $R_{\text{cryst}}$  (auch:  $R_{\text{work}}$ ):

$$R_{\text{cryst}} = \frac{\sum_{hkl} \left| |F_{\text{obs}}| - k \times |F_{\text{calc}}| \right|}{\sum_{hkl} |F_{\text{obs}}|} \quad \text{Gleichung 3-37}$$

$$\text{mit } k = \frac{w \times |F_{\text{obs}}| \times |F_{\text{calc}}|}{\sum_{hkl} w \times |F_{\text{calc}}|^2} \quad \text{Gleichung 3-38}$$

*w = auflösungsabhängiger Gewichtungsfaktor*

Zu Beginn der Verfeinerung liegt  $R_{\text{cryst}}$  häufig bei ca. 0.5, d.h. 50% (zum Vergleich: bei zufälliger Anordnung einer Polypeptidkette beträgt der  $R$ -Faktor ca. 59% [Wilson 1949]). Verfeinerte Röntgenstrukturen erreichen dagegen, je nach Auflösung der röntgenographischen Daten, kristallographische  $R$ -Faktoren von 10-30%.

Moderne Programme der Strukturverfeinerung arbeiten mittlerweile fast ausschließlich mit Hilfe von maximum-likelihood-Statistiken. Die Erläuterung der grundlegenden Prinzipien der Verfeinerung soll jedoch im Folgenden anhand der anschaulicheren *Methode der kleinsten Fehlerquadrate* (engl.: *least squares refinement*) erfolgen.

Ziel der *least squares*-Verfeinerung ist die Minimierung der folgenden Funktion:

$$Q = \sum_{hkl} w \times (|F_{\text{obs}}| - |F_{\text{calc}}|)^2 \quad \text{Gleichung 3-39}$$

mit  $w = 1/\sigma^2$   
 $\sigma$  = Standardabweichung

Das Minimum dieser Funktion erhält man durch Variation der atomaren Parameter  $u_j$ , indem man die Ableitungen der Funktion nach  $u_j$  bildet und sie gleich null setzt.

$$\frac{\partial Q}{\partial u_j} = 2 \sum_{hkl} w \times (|F_{\text{obs}}| - |F_{\text{calc}}|) \frac{\partial |F_{\text{calc}}|}{\partial u_j} = 0 \quad \text{Gleichung 3-40}$$

Ableitungen nicht-linearer Funktionen sind sehr komplex und erfordern daher erhebliche Rechnerkapazitäten. Eine wichtige Voraussetzung für die Lösung des Problems ist überdies die Verfügbarkeit ausreichender Daten. Im Falle von makromolekularen Strukturen ist das Verhältnis der *Observablen* (gemessene Reflexe) zu den *Variablen* (Atomkoordinaten und  $B$ -Faktoren) im Vergleich zu Kleinmolekülstrukturen oft ungünstig. Für die in dieser Arbeit besprochene 2.5Å-Struktur des CrotonylCoA-GcdA-Komplexes beispielsweise lag das Observablen/Variablen-Verhältnis bei ca. 1.5. Das Gleichungssystem ist also nur knapp überbestimmt (im Vergleich dazu ist das Verhältnis bei Kleinmolekülstrukturen oft > 20). Um diesem Problem zu begegnen, werden durch zusätzliche Terme in Gleichung 3-40 weitere Bedingungen in die Verfeinerung eingeführt. Die stereochemischen Parameter von Peptidstrukturen (Bindungsabstände, Bindungswinkel, Chiralität und Torsionswinkel) z.B. sind sehr genau aus den Röntgenstrukturen kleinerer Moleküle bekannt [Engh & Huber 1991]. Sie können verwendet werden, um die Werte der verfeinerten Parameter auf die bekannten Mittelwerte zu beschränken. Diese Einschränkungen (engl.: *restraints*) wirken dabei wie zusätzliche, künstliche Observablen. Benutzt man die tabellierten Werte hingegen, um die stereo-

chemischen Parameter weitgehend konstant zu halten, wird dadurch die Zahl der Variablen reduziert und man spricht von so genannten *constraints*. Analog dazu können ganze Bereiche des Modells als unveränderlich angenommen oder die Ähnlichkeit NCS-verwandter Moleküle zur Begrenzung der freien Parameter verwendet werden. Bei diesen Verfahren handelt es sich um *rigid body*- (dt.: starrer Körper) [Sussman *et al.* 1985] bzw. *NCS-Verfeinerung* [Kleywegt 1996].

Um eine Überinterpretation (engl.: *overfitting*) der gemessenen Daten bei niedrigem Daten-Parameter-Verhältnis zu verhindern, wird analog zum kristallographischen R-Faktor der *freie R-Faktor*,  $R_{free}$  [Brünger 1993; Brünger 1992] als zusätzliches Gütekriterium eingeführt:

$$R_{free} = \frac{\sum_{hkl \in T} ||F_{obs}| - k \times |F_{calc}||}{\sum_{hkl \in T} |F_{obs}|} \quad \text{Gleichung 3-41}$$

Dafür wird ein Teil der gemessenen Daten, der so genannte Testdatensatz T (in der Regel ca. 5% der einzigartigen Reflexe), nicht in die Verfeinerung der Parameter und die Berechnung der Elektronendichte einbezogen, sondern ausschließlich zur Ermittlung des freien R-Faktors verwendet. Um sicherzustellen, dass bei Vorliegen von NCS Arbeits- und Testsatz unabhängig voneinander sind, werden die Testreflexe aus einer Reihe dünner Auflösungsschalen im reziproken Raum ausgewählt.

Für die Verfeinerung wurde in Rahmen der vorliegenden Arbeit das auf einem *maximum likelihood*-Algorithmus basierende Programm *REFMAC* [Murshudov *et al.* 1997] benutzt. Ein wesentlicher Unterschied zur *least squares*-Methode liegt dabei darin, dass für die Abweichung der Observablen keine einfache Gauß-Verteilung angenommen wird, wodurch eine verbesserte Annäherung an die Realität erreicht wird. Ein weiterer Vorteil der *maximum-likelihood*-Methode ist es, dass die  $\sigma_A$ -Werte nur aus Reflexen des Testdatensatzes abgeschätzt werden, wodurch sich der model bias in  $(2mF_{obs}-DF_{calc})$ -Elektronendichtekarten verringert. Die Auswahl des Testdatensatzes wurde mit Hilfe des Programms *xdIDATAMAN* [Kleywegt & Jones 1996] getroffen.

### 3.4.14 Bioinformatische Charakterisierung der Strukturmodelle

Zusätzlich zu den bereits dargestellten Programmen zur Prozessierung (Integration, Skalierung und Reduktion) der röntgenographischen Rohdaten, der Strukturlösung und der Modellverfeinerung wurde im Rahmen dieser Arbeit eine Reihe weiterer Programme und Ser-

ver-basierter bioinformatischer Anwendungen zur Charakterisierung und graphischen Darstellung der Strukturmodelle eingesetzt. Diese sollen hier in Form einer Auflistung kurz vorgestellt werden.

### **APBS [Baker et al. 2001]**

*APBS* (Adaptive Poisson-Boltzmann Solver) ist ein Software-Paket, das auf Grundlage der Poisson-Boltzmann-Gleichung die elektrostatischen Eigenschaften von Biomolekülen berechnet. Für die Kalkulation und graphische Darstellung elektrostatischer Oberflächenpotentiale von Strukturmodellen wurde im Rahmen dieser Arbeit das APBS-Plugin für Pymol verwendet.

### **CAVER [Petrek et al. 2006]**

Der *CAVER*-Algorithmus wurde für die vollautomatische Suche nach Tunneln, Kanälen und Poren in Biomolekülen und anderen makromolekularen Strukturen entwickelt. Das Programm startet hierzu an einer zuvor vom Benutzer festgelegten Stelle innerhalb der Struktur (z.B. im aktiven Zentrum eines Enzyms) und „sucht“ mit Hilfe einer speziellen Kostenfunktion einen Weg an die Oberfläche. „Günstige“ Wegpunkte liegen dabei möglichst weit weg von den Atomen der Struktur. Mit Hilfe des *CAVER*-Plugins für Pymol [Damborsky *et al.* 2007] ist es möglich, die Resultate dieser Suche graphisch darzustellen.

### **CCP4 [CCP4 1994]**

Bei der *CCP4*-Suite (Collaborative Computational Project No. 4) handelt es sich um eine Zusammenstellung von Programmen zur röntgenographischen Proteinstrukturbestimmung, die viele der schon erwähnten und gesondert zitierten Programme beinhaltet.

### **HIC-Up [Kleywegt & Jones 1998] (<http://xray.bmc.uu.se/hicup/>)**

Die *HIC-Up*-Datenbank (Hetero-compound Information Centre – Uppsala) enthält Strukturinformationen zu einer großen Anzahl von Kleinmolekülverbindungen, die in den Strukturen der RCSB PDB vorkommen.

### **MolAxis [Yaffe et al. 2008] (<http://bioinfo3d.cs.tau.ac.il/MolAxis/>)**

Der *MolAxis*-Server bietet in Analogie zu *CAVER* die Möglichkeit, innerhalb eines Strukturmodells nach Tunneln und Kavitäten zu suchen. *MolAxis* reduziert die Suche auf zwei Dimensionen und arbeitet daher deutlich schneller als *CAVER*.

**PISA [Krissinel & Henrick 2007] ([http://www.ebi.ac.uk/msdsrv/prot\\_int/pistart.html](http://www.ebi.ac.uk/msdsrv/prot_int/pistart.html))**

Der EBI-Server *PISA* (Protein Interfaces, Surfaces and Assemblies) analysiert basierend auf den Strukturkoordinaten eines Proteins mögliche Protein-Protein-Wechselwirkungen. Der *PISA*-Algorithmus berechnet außerdem Bindungs- und Solvatationsenergien für die Bildung oligomerer Proteinkomplexe und ermöglicht unter Einbeziehung entropischer Effekte die Abschätzung der wahrscheinlichen Quartärstruktur des Zielproteins. Im Rahmen dieser Arbeit wurden für die bioinformatische Analyse der Quartärstruktur außerdem die ebenfalls über EBI zugänglichen Programme *PITA* (Protein InTerfaces and Assemblies) [Ponstingl *et al.* 2003] (<http://www.ebi.ac.uk/thornton-srv/databases/pita/>) und *PQS* (Protein Quaternary Structure) [Henrick & Thornton 1998] (<http://www.ebi.ac.uk/msd-srv/pqs/>) verwendet. Diese beiden Programme basieren rekurren auf Analysen der Kristallsymmetrie und Kristallkontakte.

**PROCHECK [Laskowski et al. 1993]**

*PROCHECK* wurde verwendet, um die stereochemische Qualität der erarbeiteten Strukturmodelle zu überprüfen. Das Programm vergleicht dabei die geometrischen Parameter der Modelle mit tabellierten Werten bekannter Strukturen der gleichen Auflösung.

**Pymol [DeLano 2002]**

*Pymol* ist ein leistungsfähiges Graphikprogramm für die Darstellung von Strukturmodellen und Elektronendichten. Für die akademische Forschung ist es lizenzfrei verfügbar. Mit *Pymol* wurden sämtliche Strukturmodell-Abbildungen der vorliegenden Arbeit erstellt. Neben einer Reihe nützlicher Funktionen zur graphischen Darstellung und Manipulation der Modelle bietet es die Möglichkeit, Render-Graphiken hoher Qualität zu erstellen. Darüber hinaus lässt sich der Funktionsumfang von *Pymol* mit Hilfe spezieller Skripts oder Plugins (dt.: Erweiterungsmodule) auf einfache Weise erweitern.

**RCSB PDB [Berman et al. 2000] (<http://www.rcsb.org/pdb/home/home.do>)**

Die *RCSB PDB* (Research Collaboratory for Structural Bioinformatics Protein Data Bank) ist eine Datenbank für 3D-Strukturdaten von Proteinen und Nukleinsäuren. Sie ist die zentrale Datenbank für alle biologischen Strukturdaten und frei zugänglich. Der Großteil dort gespeicherten Strukturmodelle und -daten wurde durch Kristallstrukturanalyse oder NMR-Spektroskopie erhalten.

**SCOP [Murzin *et al.* 1995] (<http://scop.mrc-lmb.cam.ac.uk/scop/>)**

Die SCOP-Datenbank (Structural Classification of Proteins) bietet eine detaillierte und umfassende Beschreibung der strukturellen und evolutionären Verwandtschaften von Proteinen bekannter Struktur. Die Proteine werden nach dem Grad ihrer strukturellen und evolutionären Ähnlichkeit in Familien und Superfamilien unterteilt und Klassen konservierter Faltungsmuster zugeordnet.

**SSM [Krissinel & Henrick 2004] (<http://www.ebi.ac.uk/msd-srv/ssm/cgi-bin/ssmserver>)**

Der SSM-Server (engl.: *secondary-structure matching*) des European Bioinformatik Institute (EBI, <http://www.ebi.ac.uk/>) ist eine interaktive Anwendung zum dreidimensionalen Vergleich von Proteinstrukturen. Er arbeitet auf der Grundlage von Sekundärstrukturmotiven und ermöglicht sowohl den direkten Vergleich von ausgewählten Strukturmodellen als auch die automatisierte Suche nach verwandten Strukturen auf den Servern der RCSB PDB und der SCOP-Datenbank.

**STRIDE [Heinig & Frishman 2004] (<http://webclu.bio.wzw.tum.de/stride/>)**

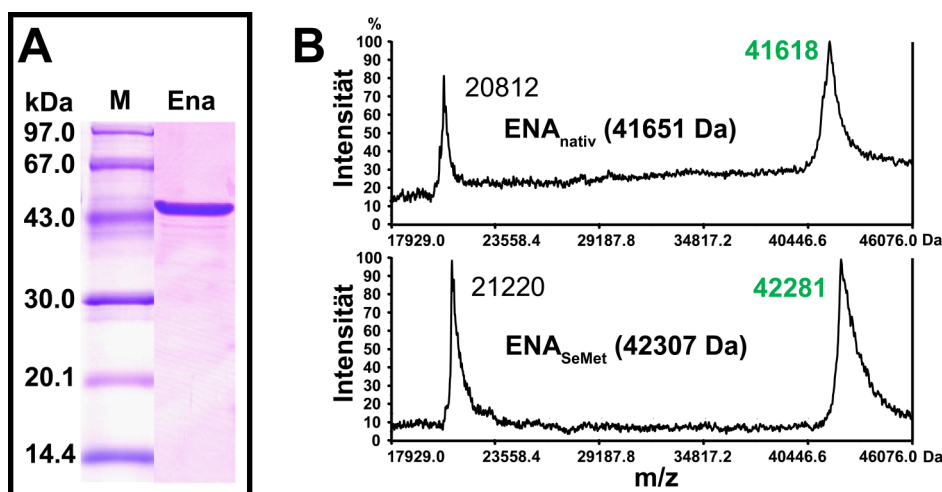
Mit Hilfe des STRIDE-Servers (secondary STRuctural IDentification) ist es möglich, die Sekundärstruktur motive von Proteinstrukturmodellen vorherzusagen. In einem wissensbasierten Algorithmus werden dabei die berechneten erwarteten Energien von Wasserstoffbrückenbindungen und statistische Informationen über die Torsionswinkel des Proteinrückgrats miteinander kombiniert. Das zugrunde liegende Programm [Frishman & Argos 1995] ist frei verfügbar und kann mit Hilfe des Skripts *stride\_ss.py* bei der Darstellung von Proteinstrukturen mit Pymol verwendet werden. Das Skript ist unter folgender Internetadresse frei erhältlich: <http://pldserver1.biochem.queensu.ca/~rlc/work/pymol/>.

## 4 Ergebnisse

### 4.1 Die Kristallstruktur der *Eubacterium barkeri*-Enamidase

#### 4.1.1 Klonierung, Expression und Aufreinigung

Die Klonierung und Expression des Enamidase-Gens aus *Eubacterium barkeri* sowie die Aufreinigung der Enamidase als Fusionsprotein mit N-terminalem Strep-tag wurden von Aschraf Alhapel und Antonio Pierik durchgeführt [Alhapel *et al.* 2006]. Verwendet wurden dabei der Plasmidvektor pPR-IBA2 und der Expressionswirt *E. coli* BL21(DE3). Vor der Kristallisation wurde das aufgereinigte Enzym einer Volumenausschlusschromatographie auf einer Superdex 200 HR 10/30-Säule (GE Healthcare) unterzogen (Puffer: 100 mM Tris-HCl, pH 8.0 und 100 mM NaCl). Es eluierte dabei als Tetramer mit einem apparenten Molekulargewicht von 178 kDa (berechnet ca. 167 kDa). Anschließend wurde das Protein auf bis zu 20 mg/mL konzentriert (Abbildung 4-1A). Die Selenomethionin-markierte Enamidase wurde

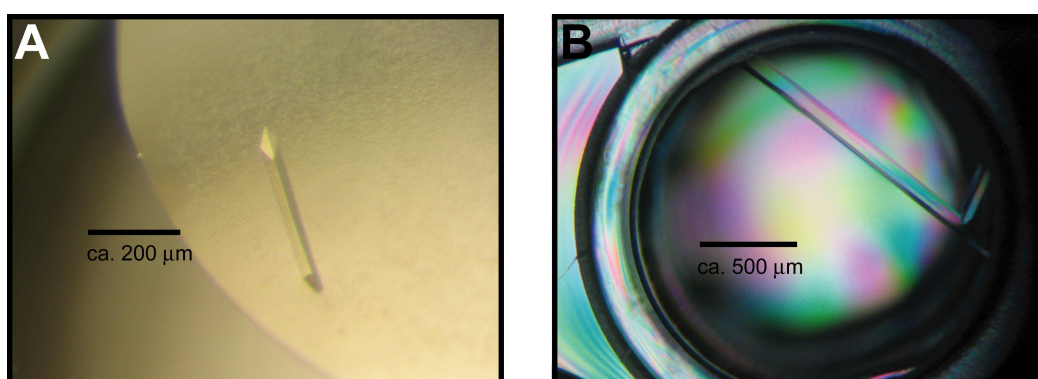


**Abbildung 4-1 Aufreinigung der Enamidase aus *E. barkeri*** (A) Coomassie-gefärbtes SDS-PAGE-Gel (12%) einer konzentrierten Enamidase-Probe nach der Gelfiltration (B) MALDI-TOF-Analyse der nativen und der SeMet-markierten Enamidase (berechnetes Molekulargewicht jeweils in Klammern). Die Differenz der experimentell ermittelten Massen (grün) entspricht dem Einbau von 14 Selenomethioninen ( $M_s = 32,07$ ,  $M_{Se} = 78,96$ ) und bestätigt damit die Vollständigkeit der SeMet-Markierung. Der zweite Peak entspricht jeweils dem doppelt geladenen Enamidase-Ion. Abbildung nach [Alhapel 2005]. Die Abweichungen von den berechneten Massen liegen im Bereich der Messungenauigkeit.

unter Anwendung des von Van Duyne et al. [Van Duyne *et al.* 1993] beschriebenen Verfahrens hergestellt. Das Aufreinigungsprotokoll entsprach dabei dem für die native Enamidase. Der Erfolg des Markierungsexperiments wurde durch MALDI-TOF-Analyse der gereinigten Enzyme überprüft (Abbildung 4-1).

#### 4.1.2 Kristallisation

Das erste Kristallisationsscreening für Enamidase wurde in 300 + 300 nL-*sitting drops* mit Hilfe eines Microsys 4004 Cartesian-Kristallisationsroboters bei Proteinkonzentrationen von 5-10 mg/mL durchgeführt. Dabei wurde eine Auswahl kommerziell erhältlicher Kristallisationslösungen eingesetzt. Die Inkubation der Ansätze erfolgte bei 18 °C. Das Enzym kristallisierte schon nach wenigen Minuten in einer großen Anzahl von Bedingungen. Allerdings zeigte sich nach einer ersten röntgenographischen Charakterisierung der Kristalle mit Hilfe der hauseigenen Cu-Drehanode, dass auf Grund einer starken Neigung zur Nukleation und Verzwillingung nur wenige Kristalle von ausreichender Qualität für Beugungsexperimente waren. Daher wurde die manuelle Optimierung Erfolg versprechender Bedingungen mittels der *hanging drop*-Methode bei 4-6 °C im Kühlraum durchgeführt, um den Prozess der Kristallisation zu verlangsamen. Die endgültige Kristallisationsbedingung für die native Enamidase setzte sich folgendermaßen zusammen: 0.1 M Tris/HCl, pH 7.0, 17.5% (w/v) PEG8000, 0.2 M Ammoniumsulfat. Kristalle des SeMet-markierten Enzyms wurden unter denselben Bedingungen in 0.1 M Cacodylat-Puffer (pH 6.5) erhalten.



**Abbildung 4-2 Kristallisation der *Eubacterium barkeri*-Enamidase** (A) Kristall der nativen Enamidase (0.1 M Tris/HCl, pH 7.0, 17.5% (w/v) PEG8000, 0.2 M Ammoniumsulfat) (B) Sehr großer Enamidase-Kristall mit der typischen säulenförmigen Morphologie (0.1 M Tris/HCl, pH 7.0, 17.5% (w/v) PEG8000, 0.2 M Ammoniumsulfat, 100 mM 2-(Hydroxymethyl)-glutarat).



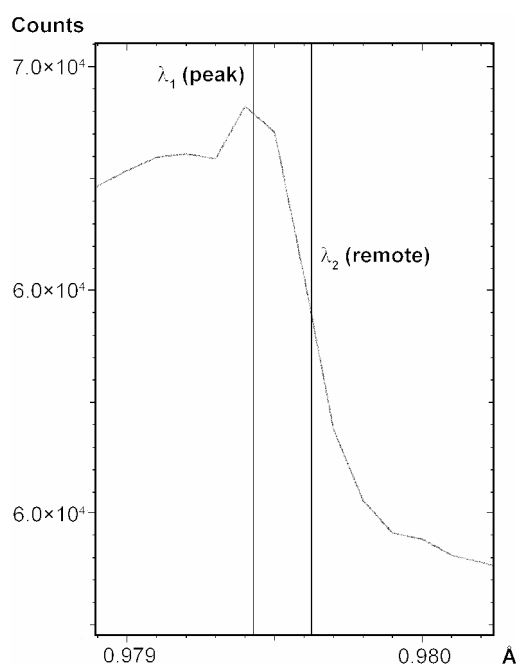
Für Kokristallisations- und *soaking*-Experimente wurden die Liganden zunächst in Gelfiltrationspuffer gelöst (Endkonzentration: 200 mM) und der pH-Wert nötigenfalls durch Zugabe geringer Mengen Natronlauge nachträglich eingestellt. Im Falle der Kokristallisation wurden die Lösungen direkt zur Proteinprobe gegeben (Endkonzentration der Liganden: 25-100 mM), während sie bei *soaking*-Experimenten direkt zur Kristallisationsbedingung hinzugefügt wurden. Um höhere Konzentrationen (bis zu 200 mM) zu erreichen, wurden die Liganden beim *soaking* zum Teil direkt in der Bedingung aufgelöst. Bei Experimenten mit dem Hydrolyse-empfindlichen Substrat THON wurde zudem zusätzlich in Gegenwart von bis zu 0.5 M  $\text{NH}_4\text{Cl}$  gearbeitet. Abbildung 4-2 zeigt beispielhaft einige Kristalle, welche die für Enamidase typische säulenförmige Morphologie aufweisen.

### 4.1.3 Datensammlung und -auswertung

Die Enamidase-Kristalle wurden vor ihrer röntgenographischen Vermessung in flüssigem Stickstoff schockgefroren. Als Kryo-Puffer diente dabei die mit 20-30% Glycerin versetzte jeweilige Kristallisationslösung. Die Datensammlung erfolgte bei 100 K am hauseigenen Röntgengenerator bzw. unter Verwendung von Synchrotronstrahlung am Deutschen Elektronen Synchrotron (DESY) in Hamburg. Die anschließende Integration und Reduktion der Messdaten wurde mit den Programmen MOSFLM und SCALA bzw. mit dem Programmpaket DENZO/SCALEPACK durchgeführt. Die jeweils ermittelten Intensitäten wurden mit dem Programm TRUNCATE in Strukturfaktoramplituden umgerechnet.

Die Vermessung von MAD-Datensätzen setzt hochenergetische Röntgenstrahlung modulierbarer Wellenlänge voraus und erfolgte daher an der MPG-ASMB-Messstation BW6 am DESY. Um die geeigneten Wellenlängen für das MAD-Experiment zu finden, wurden mit den SeMet-Kristallen zunächst Röntgenfluoreszenz-Messungen durchgeführt. In Abbildung 4-3 ist die graphische Auswertung einer solchen Messung an der Selen-Absorptionskante zur Bestimmung von peak- $(\lambda_1)$  und remote-Wellenlänge  $(\lambda_2)$  dargestellt.

Es wurden zwei Datensätze bei 0.97943  $\text{\AA}$  ( $\lambda_1$ ) bzw. 0.97963  $\text{\AA}$  ( $\lambda_2$ ) sowie ein Referenzdatensatz (remote) bei 0.95000  $\text{\AA}$  gemessen. Die maximale Auflösung nach Skalierung und Reduktion der Daten betrug 2.5  $\text{\AA}$ . Zusätzlich wurde ein 1.9  $\text{\AA}$ -Datensatz eines nativen Enamidase-Kristalls aufgenommen (nativ1). Generell war festzustellen, dass die Qualität der Kristalle starken Schwankungen unterlag. So mussten für die Vermessung des nativen Datensatzes zunächst mehr als 40 Kristalle derselben Kristallisationsbedingung getestet werden.



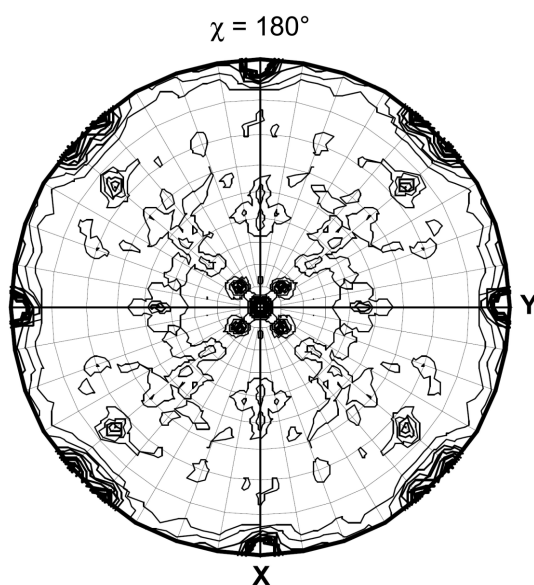
**Abbildung 4-3 Röntgenfluoreszenz-Scan zur Bestimmung der MAD-Wellenlängen für einen SeMet-Enamidase-Kristall**

#### 4.1.4 Strukturlösung und -verfeinerung

Der erste vollständige Datensatz für einen nativen Enamidase-Kristall wurde mit Hilfe der hauseigenen Cu-Drehanode bis zu einer maximalen Auflösung von  $2.0 \text{ \AA}$  gemessen (nativ2). Nach der Indizierung wurde auf Grund der systematisch absenten Reflexe für  $h+k = 2n + 1$  und  $h = 0, k = 0, l = 2n + 1$  auf die Anwesenheit einer Schraubachse und damit auf Raumgruppe  $C222_1$  geschlossen. Die Berechnung des Matthews-Koeffizienten deutete auf das Vorhandensein von 4 Molekülen in der asymmetrischen Einheit hin. Die Interpretation der mit MOLREP berechneten Selbstrotationsfunktion legte den Schluss nahe, dass die Monomere wahrscheinlich als 222-punktsymmetrisches Tetramer organisiert sind: Für  $\chi = 180$  sind Peaks für drei orthogonale 2-zählige Symmetrieachsen zu beobachten (Abbildung 4-4).

Die Lösung des Phasenproblems durch MAD-Phasierung wurde mit Hilfe von SeMet-markierten Kristallen durchgeführt. Zuvor waren Versuche gescheitert, die Struktur mittels Molekularem Ersatz unter Verwendung bekannter Amidohydrolase-Strukturen aufzuklären. Mit dem Programm SOLVE konnten 51 der insgesamt 52 Selen-Positionen bestimmt und

eine erste Elektronendichte berechnet werden. Der mittlere FOM der von SOLVE bis zu einer Auflösung von 2.5 Å bestimmten experimentellen Phasen lag bei 0.587. Sie wurden mit dem Programm RESOLVE unter Ausnutzung der nicht-kristallographischen 222-Symmetrie des Enamidase-Tetramers verfeinert und mit Hilfe des nativen 1.9 Å-Datensatzes auf 2.2 Å erweitert (FOM = 0.685). Der Modellbau und der Einbau der Solvensmoleküle konnten weitestgehend automatisch mit dem Programm ARP/wARP durchgeführt werden. Dabei wurden die experimentellen Phasen durch Kombination mit den nativen Enamidase-Daten auf 1.9 Å erweitert. Anschließend erfolgte die Verfeinerung des Strukturmodells mit den Programmen REFMAC und COOT. Der FOM für das finale Modell betrug 0.868. Eine umfassende Übersicht über die Datensammlungs-, Phasierungs- und Verfeinerungsstatistiken ist Tabelle 4-1 zu entnehmen. Die experimentellen Strukturfaktoren und die Koordinaten des Strukturmodells wurden in der RCSB PDB hinterlegt und stehen unter dem Zugangscode 2vun zur Verfügung.



**Abbildung 4-4 Selbstrotationsfunktion eines Enamidase-Datensatzes bei konstantem Winkel  $\chi = 180^\circ$**  Für die Berechnung mit Hilfe des Programms MOLREP wurden Daten von 24.8-2.5 Å verwendet.

	SeMEt			nativ1	nativ2
Datensammlung					
Raumgruppe	C2221			C2221	C2221
Zellparameter					
a (Å)	146.03			146.35	146.37
b (Å)	161.16			159.82	160.05
c (Å)	162.37			161.68	161.15
Wellenlänge (Å)	0.97943 ( $\lambda_1$ )	0.97963 ( $\lambda_2$ )	0.95000 ( $\lambda_3$ )	1.05000	1.5418
Auflösung (Å)	40.4-2.5	40.4-2.5	40.3-2.5	39-1.89	25-2.00
Vollständigkeit (%)	100 (100)	100 (100)	100 (100)	99 (91)	100 (100)
Gemessene Reflexe	243323	244157	244625	517787	1442006
Einzigartige Reflexe	65898	65951	65905	148870	126580
Multiplizität	3.7	3.7	3.7	3.2	11.4
R <sub>merge</sub> (%)	5.6 (11.6)	5.4 (11.2)	5.1 (11.9)	4.4 (25.4)	5.5 (38.8)
I/ $\sigma$	16.0 (8.4)	16.7 (8.6)	17.5 (8.8)	18.5 (2.9)	16.5 (2.3)
Phasierung					
Streufaktor f' (e-)	-6.0	-7.0	-3.0		
Streufaktor f'' (e-)	5.0	3.5	2.5		
Mittlerer FOM	0.59				
Verfeinerung					
Auflösung (Å)				38.78-1.89	
verwendete Reflexe (F>0)				146770	
R <sub>cryst</sub> (%)				17.6 (24.1)	
R <sub>free</sub> (%)				19.8 (26.4)	
Anzahl Proteinatome				12405	
Anzahl Wassermoleküle				1235	
Anzahl Metallionen				8	
Mittlerer B-factor (Å²)				25.1	
r.m.s.d. <sup>1</sup>					
Bindungslängen (Å)				0.008	
Bindungswinkel (°)				1.097	

**Tabelle 4-1 Datensammlungs-, Phasierungs- und Verfeinerungsstatistiken** Die Werte in Klammern beziehen sich auf die höchste Auflösungsschale.

<sup>1</sup> r.m.s.d  $\triangleq$  root mean square deviation (dt.: Standardabweichung)

Die Versuche, durch Kokristallisations- und *soaking*-Experimente eine Komplexstruktur von Enamidase mit einer der in 3.3.3 beschriebenen Verbindungen zu erhalten, sind leider erfolglos geblieben. Obwohl das Enzym in Gegenwart hoher Konzentrationen der Liganden problemlos Kristalle von guter Qualität bildete, lieferte die Auswertung zahlreicher Datensätze keine signifikante zusätzliche Elektronendichte im aktiven Zentrum. Es wird daher davon ausgegangen, dass Enamidase für die Verbindungen 6-Oxo-1,4,5,6-tetrahydro-pyridazin,

6-Hydroxynicotinat und 2-(Hydroxymethyl)-glutarat keine oder nur eine unzureichende Affinität besitzt. Im Falle des Substrates THON wird vermutet, dass seine Hydrolyse-Empfindlichkeit die Bildung eines in der Kristallstruktur nachweisbaren Enzym-Substrat-Komplexes verhindert [Bozarth 2005].

#### 4.1.5 Qualität des Strukturmodells

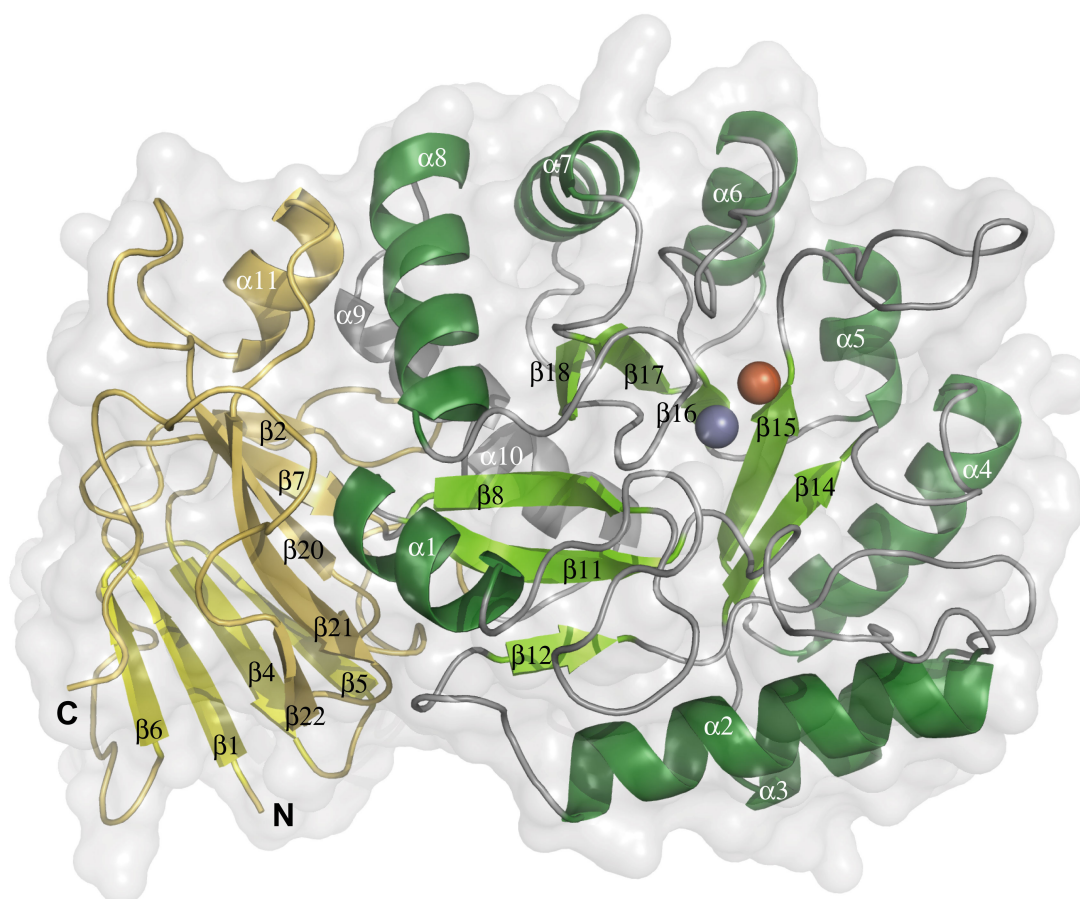
Die asymmetrische Einheit der Kristallstruktur enthält neben dem Enamidase-Tetramer jeweils 4 Eisen- und Zink-Atome, vier Chloridionen, fünf Glycerin- und 1235 Wassermoleküle. Die Aminosäurereste Ser2-Leu386 der Monomere sind sehr gut durch die Elektronendichte definiert und zeigen exzellente geometrische Eigenschaften: Die Standardabweichungen für die Bindungslängen und -winkel liegen bei 0.008 Å bzw. 1.097°; 87.3% der Aminosäurereste befinden sich in den bevorzugten Bereichen des mit dem Programm PROCHECK berechneten Ramachandran-Plots. Das finale Modell hat einen durchschnittlichen B-Faktor von 25.1 Å<sup>2</sup> und wurde bis zu einem  $R_{\text{cryst}}$  von 17.6% und einem freien R-Faktor von 19.8% verfeinert.

#### 4.1.6 Beschreibung des Enamidase-Strukturmodells

Abbildung 4-5 zeigt das Strukturmodell des Enamidase-Monomers in der Bänderdarstellung. Die Sekundärstrukturmodule wurden mit dem Programm STRIDE bestimmt und umfassen insgesamt 11  $\alpha$ -Helices ( $\alpha$ 1- $\alpha$ 11), 4  $3_{10}$ -Helices ( $\eta$ 1- $\eta$ 4) und 23  $\beta$ -Stränge ( $\beta$ 1- $\beta$ 23). In der für die Familie der Amidohydrolasen typischen Weise faltet sich das Enamidase-Monomer in zwei charakteristische Domänen: eine zentrale katalytische Domäne (Ser62-Asn321), die von einem verzerrten ( $\alpha/\beta$ )<sub>8</sub>-TIM-barrel-Strukturmotiv ( $\alpha$ 1- $\alpha$ 8,  $\beta$ 8,  $\beta$ 11- $\beta$ 18) dominiert wird, und eine zusammengesetzte  $\beta$ -Sandwich-Domäne, die sowohl aus den N-terminalen Aminosäuren Ser2-Pro61 als auch den C-terminalen Resten Thr322-Leu386 ( $\beta$ 1- $\beta$ 7,  $\beta$ 19- $\beta$ 23) besteht.

Das TIM-barrel-Strukturmotiv der Enamidase-Struktur ist im Vergleich zur Triosephosphat-Isomerase [Banner *et al.* 1976] elliptisch verzerrt: Ihre  $\beta$ -Stränge sind deutlich kürzer und bilden keine durchgängige  $\beta$ -Fass-Struktur. Das „Enamidase-Fass“ besteht vielmehr aus zwei unabhängigen parallelen  $\beta$ -Faltblättern, die insgesamt neun  $\beta$ -Stränge umfassen ( $\beta$ 8,  $\beta$ 11,  $\beta$ 12 und  $\beta$ 13- $\beta$ 18). Der Grund für diese Diskrepanz liegt darin, dass die beiden aufeinander folgenden Stränge  $\beta$ 12 und  $\beta$ 13 das ( $\alpha/\beta$ )-Motiv aufbrechen, da sie unterschiedlichen

Teilen des Fasses angehören. Zusätzlich gibt es ein kurzes antiparalleles  $\beta$ -Faltblatt ( $\beta 9$  und  $\beta 10$ ) in der Nähe des N-Terminus des Fasses zwischen  $\beta 8$  und  $\alpha 1$ . Der C-terminale Abschluss der katalytischen Domäne wird durch zwei aufeinander folgende  $\alpha$ -Helices ( $\alpha 9$  und  $\alpha 10$ ) gebildet.

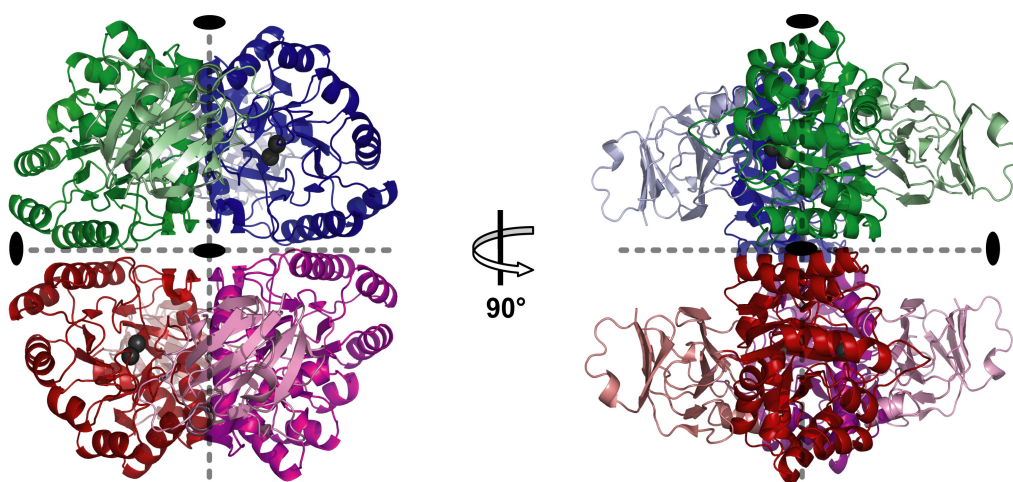


**Abbildung 4-5 Monomer-Struktur der *E. barkeri*-Enamidase** Die Sekundärstrukturmodule wurden mit Hilfe des Programms STRIDE [Heinig & Frishman 2004] bestimmt.  $3_{10}$ -Helices und  $\beta$ -Stränge kleiner als drei Aminosäurereste wurden der Übersicht halber weggelassen. Die zentrale Domäne des Amidohydrolasen-Faltungsmotivs (Gly62-Asn321) ist in grau und in Grünschattierungen dargestellt, die periphere Kompositdomäne (Ser2-Pro61, Thr322-Leu386) in hell- und dunkelgelb. Die  $\alpha$ -Helices des TIM-barrel-Strukturmotivs sind dunkelgrün, die entsprechenden  $\beta$ -Stränge in hellgrün gefärbt. Eisen- und Zinkatome im aktiven Zentrum sind als orange bzw. blaue Kugeln gezeigt.

Die zusammengesetzte  $\beta$ -Sandwich-Domäne umfasst die Aminosäurereste Ser2-Pro61 und Thr322-Leu386 und besteht aus einem fünfsträngigen ( $\beta 1$ ,  $\beta 4$ -6,  $\beta 23$ ) und einem sechssträngigen  $\beta$ -Faltblatt ( $\beta 2$ ,  $\beta 3$ ,  $\beta 7$ ,  $\beta 20$ -22). Die Stränge sind dabei bis auf die Paare  $\beta 1/\beta 6$  und  $\beta 2/\beta 7$  antiparallel. Das kleinere Faltblatt bildet den oberen Abschluss des „Sandwichs“, wohingegen sich das größere an die katalytische Domäne „schmiegt“.

### 4.1.7 Quartärstruktur

In Übereinstimmung mit den Daten der Gelfiltration liegt die *E. barkeri*-Enamidase im Kristall als 222-symmetrisches Tetramer vor. Eine Überlagerung der Proteinketten A bis D mit Hilfe des SSM-Servers ergab für die Standardabweichungen der  $C_{\alpha}$ -Atome der 385 äquivalenten Aminosäurereste Werte zwischen 0.14 (A-B) und 0.25 Å (A-C). Die Vertreter der Amidohydrolase-Familie nehmen eine Vielzahl verschiedener Quartärstrukturen an, wie z.B. Monomere (1bf6: *E. coli*-Phosphotriesterase-homologes Protein [Buchbinder *et al.* 1998]), Dimere (1j79: *E. coli*-Dihydroorotase [Thoden *et al.* 2001]), Trimere (2kau: *K. aerogenes*-Urease [Jabri *et al.* 1995]), Hexamere (1k6w: *E. coli*-Cytosin-Deaminase [Ireton *et al.* 2002]) oder sogar Oktamere (1onw: *E. coli*-Isoaspartyl-Dipeptidase [Thoden *et al.* 2003]). Interessanterweise sind die strukturell charakterisierten tetrameren Amidohydrolasen, die Dihydropyrimidinasase aus *S. kluyveri* [Lohkamp *et al.* 2006] (2fvk) und die *E. coli*-N-Acetylglucosamin-6-Phosphat-Deacetylase [Ferreira *et al.* 2006] (1yrr), in ihrer Quartärstruktur nicht verwandt mit *E. barkeri*-Enamidase.



**Abbildung 4-6 Strukturmodell des 222-symmetrischen Tetramers der *E. barkeri*-Enamidase** Gezeigt sind zwei orthogonale Orientierungen. Die Moleküle A, B, C und D sind in grün, blau, rot bzw. magenta dargestellt. Die jeweiligen peripheren Domänen sind in hellerem Farbton gehalten.

Das 222-symmetrische Enamidase-Tetramer entspricht einem Dimer von Dimeren (Abbildung 4-6). Die Monomere weisen eine durchschnittliche Oberfläche von 14,391 Å<sup>2</sup> auf und setzen sich paarweise zu den Dimeren AB und CD zusammen, wobei lösungsmittelzugängliche Oberflächen von 1,916 Å<sup>2</sup> bzw. 1,876 Å<sup>2</sup> eingeschlossen werden. Die Assoziation dieser Dimere entlang der AC- und BD-Kontaktflächen generiert zusätzliche Grenzflächen

von 877 Å<sup>2</sup> bzw. 872 Å<sup>2</sup>. Zwischen AD und BC gibt es keine Berührungsflächen. Insgesamt sind 20% der Monomerflächen an der Ausbildung des Tetramers beteiligt, wobei alle wechselwirkenden Reste dem  $\alpha/\beta$ -Fass-Teil der jeweiligen Monomere angehören. Die größere der Kontaktflächen enthält 17 Wasserstoffbrückenbindungen und eine Salzbrücke zwischen Lys79 und Asp82. Die Interaktion ergibt sich zum Teil durch Wechselwirkungen zwischen den Aminosäureresten der langen Helices  $\alpha 8$  und  $\alpha 11$  und den Schleifenregionen, die sich an die  $\beta$ -Stämme  $\beta 16$ ,  $\beta 17$  und  $\beta 19$  anschließen. Die interessanteste Wechselwirkung beinhaltet das  $\beta/\eta/\beta$ -Motiv zwischen  $\beta 8$  und  $\alpha 1$ : Strang  $\beta 10$  (z.B. in Kette A) interagiert mit den äquivalenten Resten der benachbarten Polypeptidkette ( $\beta 10'$  in Kette B), wobei ein intermolekulares antiparalleles  $\beta$ -Faltblatt der Form  $\beta 9/\beta 10/\beta 10'/\beta 9'$  entsteht. Die kleineren Kontaktflächen zwischen AC und BD kommen ausschließlich durch symmetrische Wechselwirkungen zwischen der Helix  $\alpha 2$  und der ihr vorgelagerten Schleifenregion zu Stande. Diese Wechselwirkungen werden durch sechs Wasserstoffbrückenbindungen und zwei Salzbrücken zwischen Asp109 und Lys122 etabliert.

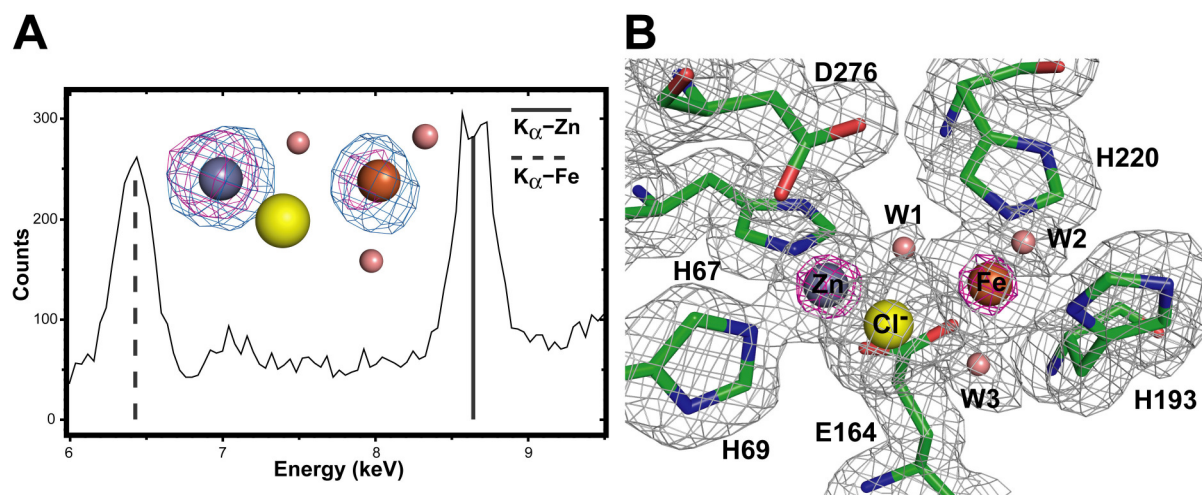
#### 4.1.8 Das aktive Zentrum der Enamidase

Wie viele Amidohydrolasen enthält auch Enamidase ein binukleares Metallzentrum in der aktiven Tasche. In den Enamidase-Kristallen wurde die Anwesenheit von Eisen ( $\lambda_{\text{Fe}} = 1.7433 \text{ Å}$ ) in den beiden Metallbindestellen  $\alpha$  und  $\beta$  durch Peaks in den anomalen Differenz-Elektronendichtekarten (Abbildung 4-7A) angezeigt, die für den mit der Cu-Drehanode ( $\lambda_{\text{CuK}\alpha} = 1.5418 \text{ Å}$ ) gemessenen nativen Datensatz (nativ2) berechnet worden waren. Das Verhältnis der anomalen Peak-Höhen an den Positionen  $\alpha$  und  $\beta$  entsprach dabei 1.2 : 1 ( $18 \pm 2 \sigma : 15 \pm 2 \sigma$ ). Allerdings zeigten Röntgenfluoreszenz-Messungen (Abbildung 4-7) signifikante Mengen von Zink in den Kristallen der rekombinanten Enamidase an. Sowohl  $\text{FeK}\alpha$ - (6.4 keV) als auch  $\text{ZnK}\alpha$ -Signale (8.6 keV) wurden detektiert, wobei die integrierten Intensitäten ein annähernd stöchiometrisches Verhältnis zueinander aufwiesen. Die Gegenwart nennenswerter Mengen an  $\text{Ni}^{2+}$  konnte ausgeschlossen werden, da kein Signal für  $\text{NiK}\alpha$  (7.5 keV) beobachtet wurde.

Die beiden Teilzentren  $\alpha$  und  $\beta$  unterscheiden sich in ihrer Koordinationschemie (Abbildung 4-7B): Das  $\alpha$ -Metallion bildet direkte Kontakte zu vier der sechs Metall-bindenden Aminosäurereste aus und ist dadurch stärker von der Proteinoberfläche eingeschlossen; das  $\beta$ -Zentrum wird nur von dreien dieser Reste koordiniert und ist daher stärker dem Lösungs-

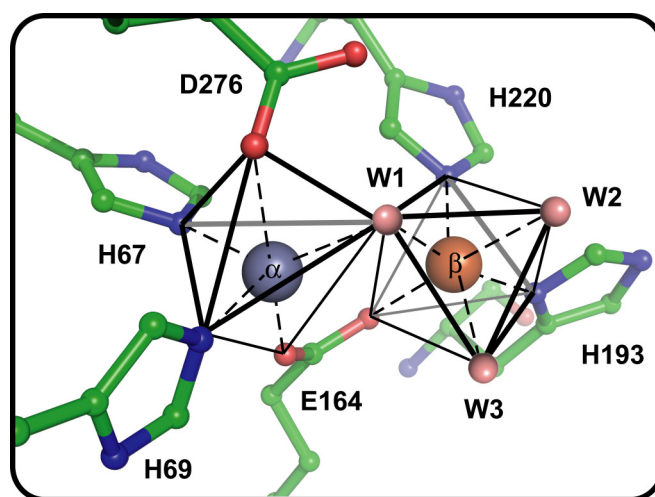


mittel ausgesetzt. Offenbar bindet Zink ( $\lambda_{\text{Zn}} = 1.2837 \text{ \AA}$ ) vorzugsweise an die  $\alpha$ -Bindestelle: Das anomale Signal des bei einer Wellenlänge von  $1.05 \text{ \AA}$  aufgenommenen Datensatzes (nativ1) war an dieser Stelle mit  $11\sigma$  ca. 1.6-mal größer als an der  $\beta$ -Bindestelle ( $6.6\sigma$ ) (Abbildung 4-7A). Um die Anwesenheit von Zink zu bestätigen und das Eisen-zu-Zink-Verhältnis zu berechnen, wurde der Metall-Gehalt zusätzlich colorimetrisch bestimmt [Alhapel 2005]. Dabei wurde der Ferrozin-ähnliche Komplexbildner Feren benutzt [Pierik *et al.* 1992], um Eisen zu bestimmen. Zink wurde mit Hilfe von 4-(2-Pyridylazo)-resorcinol nachgewiesen. Es ergab sich ein Eisen-Zink-Verhältnis von 1.1 : 0.8 - 1.0. Bemerkenswerterweise zeigten UV/VIS-Spektren der Enamidase keine signifikante Absorption; allerdings waren hochkonzentrierte Enamidase-Lösungen (20-40 mg/mL) schwach gelb gefärbt.



**Abbildung 4-7 Das binukleare Metallzentrum der Enamidase** (A) Anomale Differenzelektronendichten und Röntgenfluoreszenzspektrum. Die anomalen Differenzelektronendichten sind für die Metallionen des aktiven Zentrums gezeigt. Die zu Grunde liegenden Daten wurden mit Hilfe von Röntgenstrahlung der Wellenlänge  $1.5418 \text{ \AA}$  (blau, Konturierungsgrad  $5\sigma$ ) bzw.  $1.0500 \text{ \AA}$  (magenta, Konturierungsgrad  $5\sigma$ ) gemessen. Das Röntgenfluoreszenzspektrum wurde für einen Enamidase-Kristall mit einer Anregungsenergie von  $13.6 \text{ keV}$  aufgenommen. (B) Finales Strukturmodell des aktiven Zentrums zusammen mit der entsprechenden  $2F_{\text{obs}} - F_{\text{calc}}$ -Elektronendichte bei zwei verschiedenen Konturierungsgraden:  $1\sigma$  (grau) und  $8\sigma$  (magenta). Die Zink- und Eisenionen werden von vier Histidin-Seitenketten, einem Aspartat-Rest und drei Wassermolekülen (rot) koordiniert. Lys164 und W1 verbrücken das binukleare Metallzentrum. Zusätzliche Elektronendichte wurde einem Chloridion zugeordnet (gelb).

Vier Histidin-, ein Glutamat- und ein Aspartat-Rest koordinieren die beiden katalytischen Metallionen des aktiven Zentrums. Wie für  $(\alpha/\beta)_8$ -Fass-Enzyme typisch, sind diese Reste an den C-terminalen Enden der  $\beta$ -Stränge ( $\beta_8$ ,  $\beta_{14}$ ,  $\beta_{15}$ ,  $\beta_{16}$ ,  $\beta_{18}$ ) lokalisiert, die sich am Rand einer trichterartigen Vertiefung an einem Ende des Fasses befinden (Abbildung 4-5). Alle Reste des aktiven Zentrums einschließlich der Metallionen sind klar durch Elektronendichte definiert (Abbildung 4-7B). Der Abstand zwischen Eisen und Zink liegt bei 3.6-3.7 Å und entspricht damit den typischen Werten für Amidohydrolasen mit binuklearen Metallzentren [Seibert & Raushel 2005].



**Abbildung 4-8 Koordinationsgeometrie der Metallionen im aktiven Zentrum** Zink (blau) ist in verzerrter trigonal-bipyramidaler Weise koordiniert, Eisen (rot) verzerrt oktaedrisch.

Abbildung 4-8 illustriert die Koordinationsgeometrie des Metallzentrums. Das Zinkion an Position  $\alpha$  wird jeweils von den  $N_{\epsilon 2}$  der beiden Histidin-Reste (His67/69) eines HXH -Motivs koordiniert, das innerhalb der Amidohydrolase-Familie gut konserviert ist. Seitenketten-Sauerstoffatome von Asp276 ( $O_{\delta 2}$ ) und Glu164 ( $O_{\epsilon 2}$ ) sowie ein Wassermolekül (W1) vervollständigen die verzerrt trigonal-bipyramidale Architektur des gebundenen Zinkions. Glu164 und W1 verbrücken Letzteres mit dem Eisenion in der  $\beta$ -Position des binuklearen Metallzentrums. Auf Grund der Nähe der positiv geladenen Metallionen (2.06 und 2.25 Å) liegt das verbrückende Wassermolekül W1 wahrscheinlich deprotoniert vor und ist daher ein möglicher Kandidat für das aktive Nukleophil in der hydrolytischen Enamidase-Reaktion. Das  $\beta$ -Metall wird verzerrt oktaedrisch von His193 ( $N_{\delta 1}$ ), His220 ( $N_{\epsilon 2}$ ) und Glu164 ( $O_{\epsilon 1}$ ) sowie W1

und zwei weiteren Solvensmolekülen W2 und W3 koordiniert. Dabei besetzen W2 und der O<sub>ε1</sub>-Sauerstoff von Glu164 die apikalen Positionen. Die Metallliganden-Abstände innerhalb des binuklearen Metallzentrums liegen zwischen 2.0 und 2.5 Å. Zusätzliche Elektronendichte im Abstand von 3.2 Å von der α-Position wurde einem Chloridion zugeordnet. Es besetzt eine der theoretischen axialen Positionen eines windschiefen Oktaeders um das Zinkatom. Die Geometrie der Metallliganden entspricht interessanterweise der Koordination des binuklearen Metallzentrums der mit Cadmium rekonstituierten Phosphotriesterase [Benning *et al.* 1995].

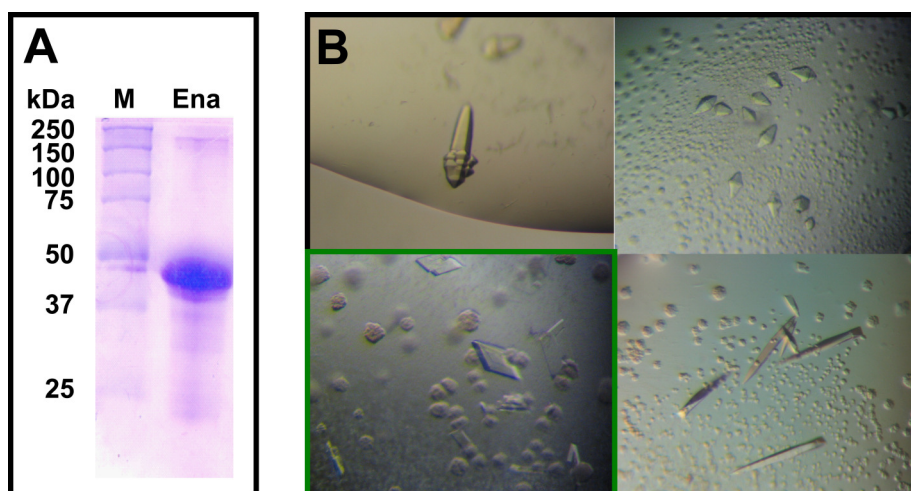


## 4.2 Die Kristallstruktur der *Bradyrhizobium japonicum*-Enamidase

In einem weiteren Kooperationsprojekt mit Andrew Bozarth und Antonio Pierik wurde die Enamidase aus dem Organismus *Bradyrhizobium japonicum*, einem Stickstoff-fixierenden Symbionten der Sojabohne (*Glycine max*) kristallisiert. Ein vorläufiges Modell der Kristallstruktur wurde bereits ermittelt, befand sich allerdings bei Abschluss dieser Arbeit noch in einem frühen Stadium der Verfeinerung. Im Folgenden kann deshalb nur auf die zu diesem Zeitpunkt bereits vorliegenden Ergebnisse der vorläufigen strukturellen Charakterisierung eingegangen werden.

### 4.2.1 Klonierung, Expression und Aufreinigung

Die Klonierung und Expression des Enamidase-Gens aus *Eubacterium barkeri* sowie die Aufreinigung und biochemische Charakterisierung der Enamidase aus *Bradyrhizobium japonicum* erfolgten im Wesentlichen analog zum *E. barkeri*-Enzym und wurden von Andrew Bozarth



**Abbildung 4-9 *Bradyrhizobium japonicum*-Enamidase** (A) Coomassiegefärbtes SDS-PAGE-Gel (12%) einer konzentrierten Enamidase-Probe nach der Umpufferung (ca. 20 µg). (B) Auswahl der erhaltenen Kristalle: links oben, 0.1 M Cacodylat-Puffer (pH 6.5), 0.2 M Ammoniumsulfat, 15% (w/v) PEG4000; rechts oben, 0.1 M HEPES (pH 7.5), 0.1 M Ammoniumacetat, 15% (w/v) PEG4000; links unten, 0.1 M Cacodylat-Puffer (pH 6.5), 0.1 M Ammoniumsulfat, 15% (w/v) PEG4000; rechts unten, 0.1 M HEPES (pH 7.5), 0.2 M Ammoniumacetat, 20% (w/v) PEG4000. Die Kristallform, von welcher der in der Strukturlösung verwendete Datensatz erhalten wurde, ist grün markiert.

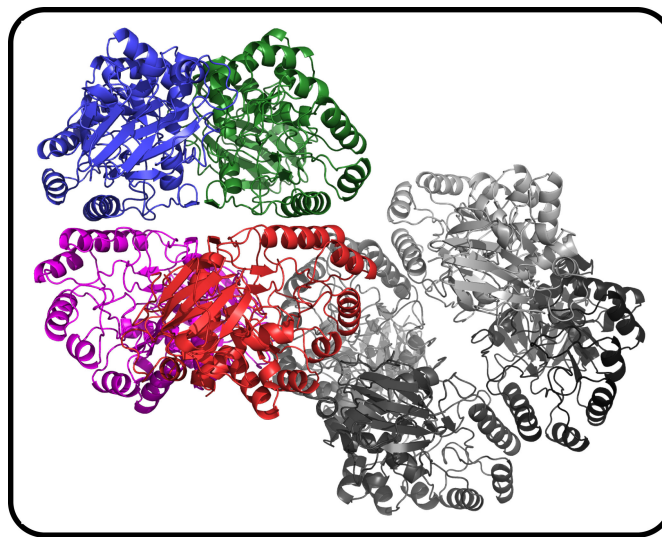
und Antonio Pierik durchgeführt [Bozarth 2005]. Für die Hydrolyse von THON wurde ein  $K_m$  von 6 mM und eine im Vergleich zur *E. barkeri*-Enamidase fast fünffach höhere  $V_{max}$  von 260 Units/mg ermittelt. Mit Hilfe der bereits erwähnten Methoden (vgl. 4.1.8) wurde ein Eisen- und Zink-Gehalt von 1.28 bzw. 0.87 pro Enamidase-Molekül bestimmt. Vor der Kristallisation wurde das aufgereinigte Enzym auf einer Econo-Pac® 10 DG-Säule (BioRad) umgepuffert (10 mM Tris, 100 mM NaCl, pH 8.0) und bis zu einer Konzentration von 10-20 mg/mL aufkonzentriert (Abbildung 4-9A).

#### 4.2.2 Kristallisation, Datensammlung und Strukturlösung

Das Kristallisationsscreening für die *B. japonicum*-Enamidase erfolgte im Rahmen des internationalen Workshops „Protein crystallization – Present and future“ an der EMBL-Niederlassung in Hamburg. Das Enzym wurde in die dortige Anlage für Hochdurchsatzkristallographie eingespeist, wobei eine Vielzahl kommerzieller Kristallisationsscreens – insgesamt 1044, z.T. allerdings redundante Bedingungen – zum Einsatz kamen. Es zeigte sich, dass das Enzym ähnlich wie die Enamidase aus *E. barkeri* bereitwillig in einer großen Anzahl von Bedingungen kristallisiert, allerdings deutlich stärker zur Ausbildung verwachsener Kristalle tendiert. Dies veranschaulicht die in Abbildung 4-9B dargestellte Auswahl von Kristallen, die bei der anschließend in Marburg durchgeführten manuellen Optimierung ausgewählter Bedingungen erhalten wurden. Die Kristalle der *Bradyrhizobium japonicum*-Enamidase zeigten dabei einen ausgeprägten Polymorphismus, wobei selbst in ein und derselben Bedingung verschiedene Kristallformen unterschiedlicher Qualität beobachtet werden konnten. Entsprechend mussten bei der röntgenographischen Vermessung an der EMBL-Messstation BW7a ( $\lambda = 0.806 \text{ \AA}$ ) am DESY (Hamburg) mehr als 40 Kristalle getestet werden, bevor ein für die Strukturlösung geeigneter Datensatz erhalten werden konnte.

Dieser Datensatz hatte eine maximale Auflösung von  $2.1 \text{ \AA}$  und wurde in Raumgruppe P1 ( $a = 79.3912 \text{ \AA}$ ,  $b = 86.7573 \text{ \AA}$ ,  $c = 136.2601 \text{ \AA}$ ,  $\alpha = 100.5703^\circ$ ,  $\beta = 91.3497^\circ$ ,  $\gamma = 115.2553^\circ$ ) prozessiert ( $R_{merge} = 26.1\%$ ). Integration und Datenreduktion wurden mit den Programmen MOLREP bzw. SCALA/TRUNCATE durchgeführt. Für die Strukturlösung durch Molekularen Ersatz wurde als Referenzstruktur das Strukturmodell der *E. barkeri*-Enamidase verwendet. Die Verfeinerung des daraus erhaltenen Modells ist noch nicht abgeschlossen, derzeit weist es ein Verhältnis von  $R_{cryst}/R_{free}$  von 19.5%/25.0% auf.

Die asymmetrische Einheit des Modells für *Bradyrhizobium japonicum*-Enamidase enthält insgesamt acht Moleküle, die in Form zweier identischer Tetramere angeordnet sind (Abbildung 4-10). In den einzelnen Monomeren konnte wie erwartet zusätzliche Elektronendichte den Metallionen des aktiven Zentrums zugeordnet werden. Die vorläufige Prüfung der Kristallstruktur bestätigt damit bereits wesentliche Teile der Analyse des *E. barkeri*-Modells. Darüber hinaus sind schon jetzt einige signifikante Abweichungen beobachtet worden, auf die in 5.1.4 kurz eingegangen werden wird.



**Abbildung 4-10** Die asymmetrische Einheit der Kristallstruktur der *Bradyrhizobium japonicum*-Enamidase. Sie enthält zwei identische Enamidase-Tetramere.

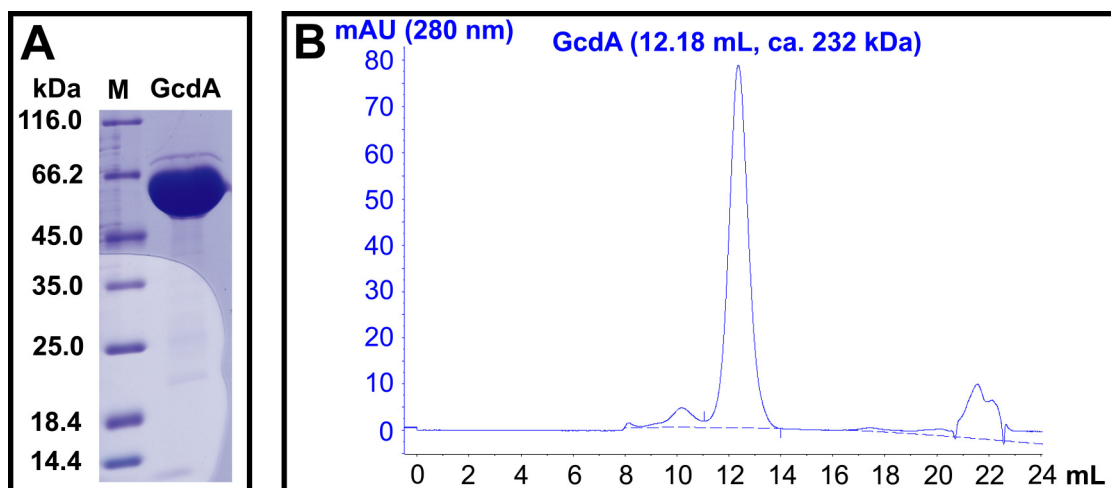




## 4.3 Die Kristallstruktur der *Clostridium symbiosum*-GcdA

### 4.3.1 Klonierung, Expression und Aufreinigung

Das *gcdA*-Gen aus *Clostridium symbiosum* wurde als Strep-tag II-Fusionsprotein in *E. coli* Rosetta™ (DE3)pLysS-Zellen (Novagen) exprimiert und über Affinitätschromatographie und anschließende Gelfiltration aufgereinigt (Abbildung 4-11). Das gereinigte Enzym eluierte dabei bei einem apparenten Molekulargewicht von 232 kDa (berechnetes Molekulargewicht des Monomers ca. 65 kDa). Nach der Auswertung des ersten Kristallisationscreenings wurde dazu übergegangen, das Protein nach der Aufreinigung in 10 mM MES (pH 6.5), 100 mM NaCl umzupuffern. In diesem Puffer konnte die GcdA problemlos auf 20-30 mg/mL aufkonzentriert werden.



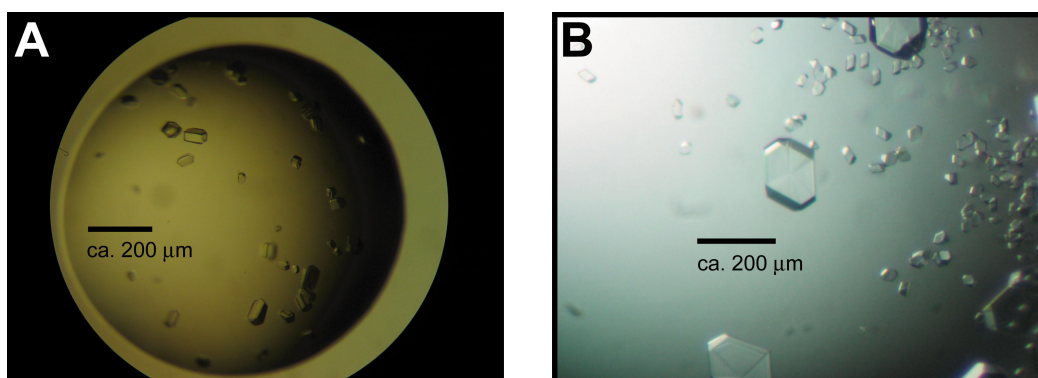
**Abbildung 4-11 Aufreinigung der GcdA aus *C. symbiosum*** (A) Coomassie-gefärbtes SDS-PAGE-Gel (12%) einer konzentrierten GcdA-Probe (30 µg) nach der Aufreinigung. (B) Chromatogramm eines typischen Gelfiltrationslaufes zur Aufreinigung und Molekulargewichtsbestimmung der GcdA (Säule: Superdex 200 HR 10/30 (GE Healthcare); Puffer: 10 mM MES, pH 6.5, 100 mM NaCl; Flussrate: 0.5 mL/min).

### 4.3.2 Kristallisation

Alle Kristallisationsexperimente wurden mit Hilfe der Dampfdiffusionsmethode bei 18 °C durchgeführt. GcdA wurde dabei in Endkonzentrationen von 5-10 mg/mL eingesetzt. Für Kokristallisationsversuche wurden 100-200 mM Lösungen von Biocytin, Crotonyl-CoA, Glutaryl-CoA und Glutaconyl-CoA in 10 mM MES (pH 6.5), 100 mM NaCl angesetzt. Diese Lö-

sungen wurden jeweils den Proteinproben vor der Kristallisation hinzugefügt. Für *soaking*-Experimente wurden die Substratlösungen direkt zu den jeweiligen Kristallisationsbedingungen gegeben. Kokristallisation und *soaking* erfolgten bei Endkonzentrationen der Liganden von 5-10 mM (CoA-Ester) und 5-50 mM (Biocytin). Ein erstes Kristallisationsscreening für GcdA wurde zunächst ohne Zugabe der Liganden (*Apo-GcdA*) in 500 + 500 nL-*sitting drops* mit Hilfe eines Microsys 4004 Cartesian-Kristallisationsroboters durchgeführt. Dabei kam eine Auswahl kommerziell erhältlicher Kristallisationslösungen zum Einsatz. Anschließend wurden Erfolg versprechende Bedingungen manuell mittels der *hanging drop*-Technik optimiert.

Die Kristalle der Apo-GcdA wurden in 0.1 M MES-Puffer, pH 6.5, 1.3-1.6 M MgSO<sub>4</sub>, 5-25% Glycerin (*Bedingung A*; Abbildung 4-12A) erhalten. Kokristalle in Gegenwart von Crotonyl-CoA wuchsen unter den gleichen Bedingungen. Obwohl sowohl von Apo-GcdA als auch von einem Crotonyl-CoA-GcdA-Kokristall vollständige Datensätze gemessen werden konnten, war die Reproduzierbarkeit dieser Kristalle schlecht, und ihre Qualität machte die Benutzung von Synchrotron-Strahlung unabdingbar. Daher wurde das Roboter-gestützte Screening für die Kokristallisation mit Glutaryl-CoA und Glutaconyl-CoA wiederholt, wodurch zwei zusätzliche Kristallisationsbedingungen gefunden wurden. Glutaryl-CoA-Kokristalle wurden dabei in 0.1 M MES, pH 6.5, 10-11% Polyethylenglycol 20000, 15-20% Glycerin (*Bedingung B*) erhalten. Die GcdA-Kristalle mit der besten Qualität wuchsen in Gegenwart von Glutaconyl-CoA in 0.1 M Natriumcitrat, pH 5.6-6.5, 1 M Amoniumphosphat, 27.5% Glycerin (*Bedingung C*; Abbildung 4-12B). Interessanterweise lieferte die letzte Bedin-

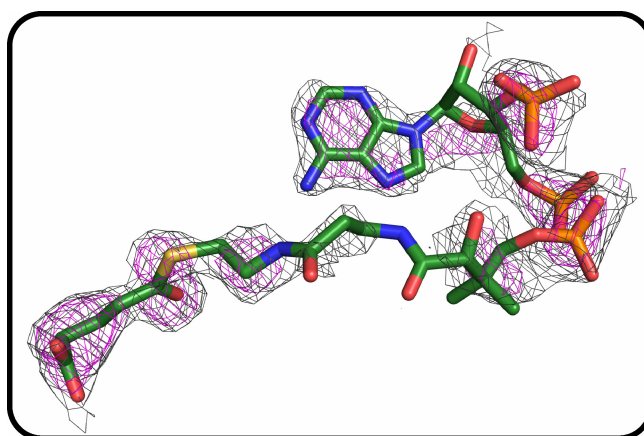


**Abbildung 4-12 Kristallisation der GcdA** (A) Kristalle von Apo-GcdA in Bedingung A (0.1 M MES-Puffer, pH 6.5, 1.3-1.6 M Magnesiumsulfat, 5-25% Glycerin) (B) Kokristallisation von GcdA mit Glutaconyl-CoA in Bedingung C (0.1 M Natriumcitrat, pH 5.6-6.5, 1 M Amoniumphosphat, 27.5% Glycerin)

gung nur Kokristalle mit Glutaconyl-CoA. Die Zugabe von Crotonyl- oder Glutaryl-CoA dagegen verhinderte entweder das Kristallwachstum oder führte zu Kristallen von deutlich schlechterer Qualität, die keines der Substrate gebunden hatten.

### 4.3.3 Datensammlung und -auswertung

Die GcdA-Kristalle wurden in ihren jeweiligen Kristallisationslösungen schockgefroren, wobei nach Bedarf zuvor 20-30% Glycerin als Kryoprotektant hinzugefügt worden waren. Die Vermessung der Kristalle erfolgte entweder an der Messstation X13 des Deutschen Elektronen Synchrotron (DESY, Hamburg) oder an der hauseigenen Cu-Drehanode ( $\lambda = 1.5418 \text{ \AA}$ ). Die maximale Auflösung der Datensätze lag dabei zwischen 2.7 und 1.75  $\text{\AA}$ . Einige der Liganden-Komplexe wurden darüber hinaus an der ESRF (Messstationen ID14-1 und ID23-2) mit zum Teil höheren maximalen Auflösungen vermessen. Allerdings zeigte sich hierbei, dass die Intensität der eingesetzten Strahlung offensichtlich zu einer Fragmentierung der Substrate führte, welche die Daten unbrauchbar machte (Abbildung 4-13).



**Abbildung 4-13 Fragmentierte Elektronendichte für Glutaryl-CoA** Dargestellt ist die  $2F_{\text{obs}} - F_{\text{calc}}$ -Elektronendichte bei einem Konturierungsgrad von  $1 \sigma$  (magenta) bzw.  $0.5 \sigma$  (grau). Der zu Grunde liegende Datensatz hatte eine maximale Auflösung von 2.1  $\text{\AA}$  und wurde an der ESRF gemessen.

Die Indizierung und Integration der Daten wurde mit dem Programm MOSFLM vorgenommen. Anschließend wurden die Datensätze jeweils mit den Programmen SCALA und TRUNCATE skaliert und reduziert. Eine Zusammenfassung der Datensammlungsstatistiken ist in Tabelle 4-2 enthalten.

Datensammlung	Apo <sup>1</sup>	Crotonyl-CoA <sup>1</sup>	Glutaryl-CoA <sup>2</sup>	Glutaconyl-CoA <sup>2</sup>
Wellenlänge (Å)	0.80150	0.80150	1.54179	1.54179
Maximale Auflösung (Å)	2.40	2.50	2.70	1.75
Höchste Auflösungsschale	2.53 - 2.40	2.64 – 2.50	2.85 - 2.59	1.84 – 1.75
Gemessene Reflexe	148483	212893	402828	239351
Einzigartige Reflexe (> 1 $\sigma$ )	24154	79234	61822	60672
Vollständigkeit (%)	99.9 (100.0)	99.6 (98.9)	98.7 (91.4)	97.9 (99.7)
I/ $\sigma$	20.9 (5.1)	9.6 (2.1)	16.6 (3.4)	18.9 (4.5)
R <sub>merge</sub>	0.074 (0.334)	0.097 (0.464)	0.103 (0.455)	0.046 (0.262)
Multiplizität	6.1 (6.2)	2.7 (2.4)	6.5 (6.1)	3.9 (3.7)
Raumgruppe	F222	C2	P3 <sub>1</sub> 21	F222
Zellparameter				
a (Å)	102.05	101.50	156.47	102.360
b (Å)	143.53	167.45	156.47	144.360
c (Å)	167.52	138.11	141.40	167.150
$\alpha$ (°)	90.00	90.00	90.00	90.00
$\beta$ (°)	90.00	89.98	90.00	90.00
$\gamma$ (°)	90.00	90.00	120.00	90.00
Moleküle in der AE <sup>3</sup>	1	4	2	1

**Tabelle 4-2 Datensammlungsstatistiken für *Clostridium symbiosum* GcdA und die entsprechenden Komplex-Strukturen** Die Werte in Klammern beziehen sich auf die höchste Auflösungsschale.

<sup>1</sup> X13 Synchrotron-Messstation (DESY, Hamburg)

<sup>2</sup> Cu-Drehanode

<sup>3</sup> AE  $\triangleq$  asymmetrische Einheit

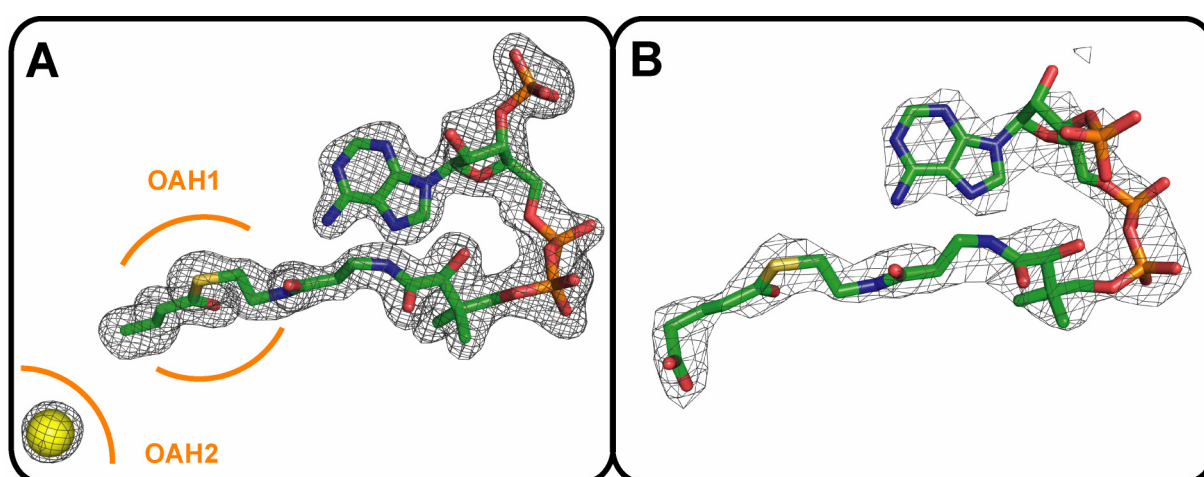
#### 4.3.4 Strukturlösung und -verfeinerung

Das Strukturmodell für Apo-GcdA wurde durch Molekularen Ersatz mit Hilfe des Programms MOLREP ermittelt. Als Referenz diente die Proteinkette A des Strukturmodells der *A. fermentans* GcdA (1pix; [Wendt *et al.* 2003]). Dabei wurden Solvensmoleküle und aus dem Suchmodell herausragende Aminosäurebereiche (Gly219-Ala253 und Thr489-Arg528) zuvor entfernt und die experimentellen Daten auf eine maximale Auflösung von 3 Å beschränkt. Anschließend wurde das vorläufige Modell durch die alternierende Verwendung der Programme COOT [Emsley & Cowtan 2004] und REFMAC [Murshudov *et al.* 1997] vervollständigt. Die Positionierung und Validierung der Solvensmoleküle erfolgte mit Hilfe der entsprechenden automatisierten Funktion von COOT und durch manuellen Einbau in die Elektronendichte. Die Strukturlösung und Verfeinerung der Liganden-Komplexe fand in analoger Weise mit Hilfe des Strukturmodells der Apo-GcdA statt.

Die Koordinaten und Strukturaktoren der verfeinerten Strukturmodelle wurden unter den folgenden Codes in der RCSB PDB hinterlegt: 3gf7 (Apo-GcdA), 3gf3 (Glutaconyl-CoA), 3glm (Crotonyl-CoA) und 3gma (Glutaryl-CoA).

### 4.3.5 Qualität der Strukturmodelle

GcdA-Kristalle von ausreichender Qualität für die röntgenographische Strukturlösung wurden in drei verschiedenen Raumgruppen und drei unterschiedlichen Kristallisationsbedingungen erhalten (vgl. 4.3.2). Die Apo-GcdA-Kristalle aus Bedingung A gehören der Raumgruppe F222 an. Das fertige Modell enthält ein Molekül in der asymmetrischen Einheit und wies nach der Verfeinerung einen  $R_{\text{cryst}}$  von 18.9% und einen freien R-Faktor von 22.7% bei einer maximalen Auflösung von 2.4 Å auf. Kokristalle mit Crotonyl-CoA wurden unter denselben Bedingungen erhalten. Bemerkenswerterweise änderte sich die Raumgruppe in Gegenwart von Crotonyl-CoA bei nahezu identischer Kristallpackung zu C2 mit vier Molekülen in der asymmetrischen Einheit. Die finalen  $R_{\text{cryst}}$ - und  $R_{\text{free}}$ -Werte dieser Struktur lagen bei 20.0% bzw. 23.8% bei einer Auflösung von 2.5 Å. Kokristalle mit Glutaryl-CoA (Bedingung B) wuchsen in Raumgruppe P3<sub>1</sub>21. Das resultierende Strukturmodell enthält zwei Moleküle in der asymmetrischen Einheit, und die Verfeinerung resultierte in einem  $R_{\text{cryst}}/R_{\text{free}}$ -Verhältnis von 19.2%/21.8% (2.7 Å). Die hochwertigsten Kristalle wurden jedoch in Bedingung C unter Zugabe von Glutaconyl-CoA erhalten. Diese Kristallform gehört ebenfalls der Raumgruppe F222 an und enthält ein Molekül in der asymmetrischen Einheit. Die finalen Werte für  $R_{\text{cryst}}$  und  $R_{\text{free}}$  lagen hier bei 16.4% bzw. 17.8% (1.75 Å). Die Substrat-Moleküle der Komplex-Strukturen sind klar durch Elektronendichte definiert (Abbildung 4-14). Allerdings



**Abbildung 4-14** (A) Verfeinertes Strukturmodell von Crotonyl-CoA in der finalen  $F_{\text{obs}}-F_{\text{calc}}$ -Elektronendichte (1.75 Å, 3  $\sigma$ ). Zusätzliche Elektronendichte in der Nähe der wahrscheinlichen Biotin-Bindestelle wurde einem Chloridion zugeordnet (gelbe Kugel). (B)  $2F_{\text{obs}}-F_{\text{calc}}$ -Elektronendichte des Substratanalogons Glutaryl-CoA (2.7 Å, 1  $\sigma$ ).

konnte im Strukturmodell der in Gegenwart von Glutaconyl-CoA gewachsenen Kristalle nur das Produkt der Decarboxylierung beobachtet werden, obwohl alle Kristalle vor der Datensammlung in Kryo-Puffern eingefroren worden waren, die die frisch angesetzten CoA-Ester enthielten.

Die Evaluierung der Strukturmodelle wurde mit dem Programm PROCHECK durchgeführt. Alle vier Strukturen zeigten gute stereochemische Eigenschaften: Mehr als 98% der Aminosäurereste befanden sich in den bevorzugten und erlaubten Regionen des Ramachandran-Plots, und die Standardabweichung der Bindungslängen und -winkel lag jeweils unterhalb von 0.010 Å bzw. 1.3°. Eine umfassende Übersicht über die Verfeinerungsstatistiken ist Tabelle 4-3 zu entnehmen. Falls nicht anders angegeben, wurde im Folgenden jeweils die am höchsten aufgelöste Struktur aus Bedingung C (Glutaconyl-CoA-Kokristallisation) für die Beschreibung und Interpretation der GcdA-Struktur verwendet.

Verfeinerung	Apo	Crotonyl-CoA	Glutaryl-CoA	Glutaconyl-CoA
R <sub>cryst</sub> /R <sub>free</sub>	0.185 / 0.225	0.194 / 0.230	0.189 / 0.210	0.163 / 0.176
r.m.s.d. Bindungslängen (Å) <sup>1</sup>	0.007	0.009	0.008	0.008
r.m.s.d. Bindungswinkel (°) <sup>1</sup>	1.095	1.202	1.155	1.173
Mittlerer B-Faktor (Å <sup>2</sup> )	31.6	35.7	36.9	23.2
Anzahl Proteinatome	4202	17298	8723	4371
Anzahl Ligandenatome	-	212	112	53
Anzahl Ionenatome	5	4	-	1
Anzahl Wassermoleküle	201	627	352	466
Ramachandran-Plot (%)				
bevorzugt	90.1	89.6	90.6	91.3
erlaubt	9.3	9.3	8.6	7.8
zusätzlich erlaubt	0.2	0.6	0.4	0.4
unerlaubt	0.4	0.4	0.4	0.4
Lücken <sup>2</sup>	G218-K251 L502-Q513	G221-A237 (C) K501-Q513 (C)	A220-A236 (A) K501-Q513 (A)	G221-A237 A505-Q513

**Tabelle 4-3 Verfeinerungsstatistiken für *C. symbiosum*-GcdA und die entsprechenden Komplex-Strukturen**

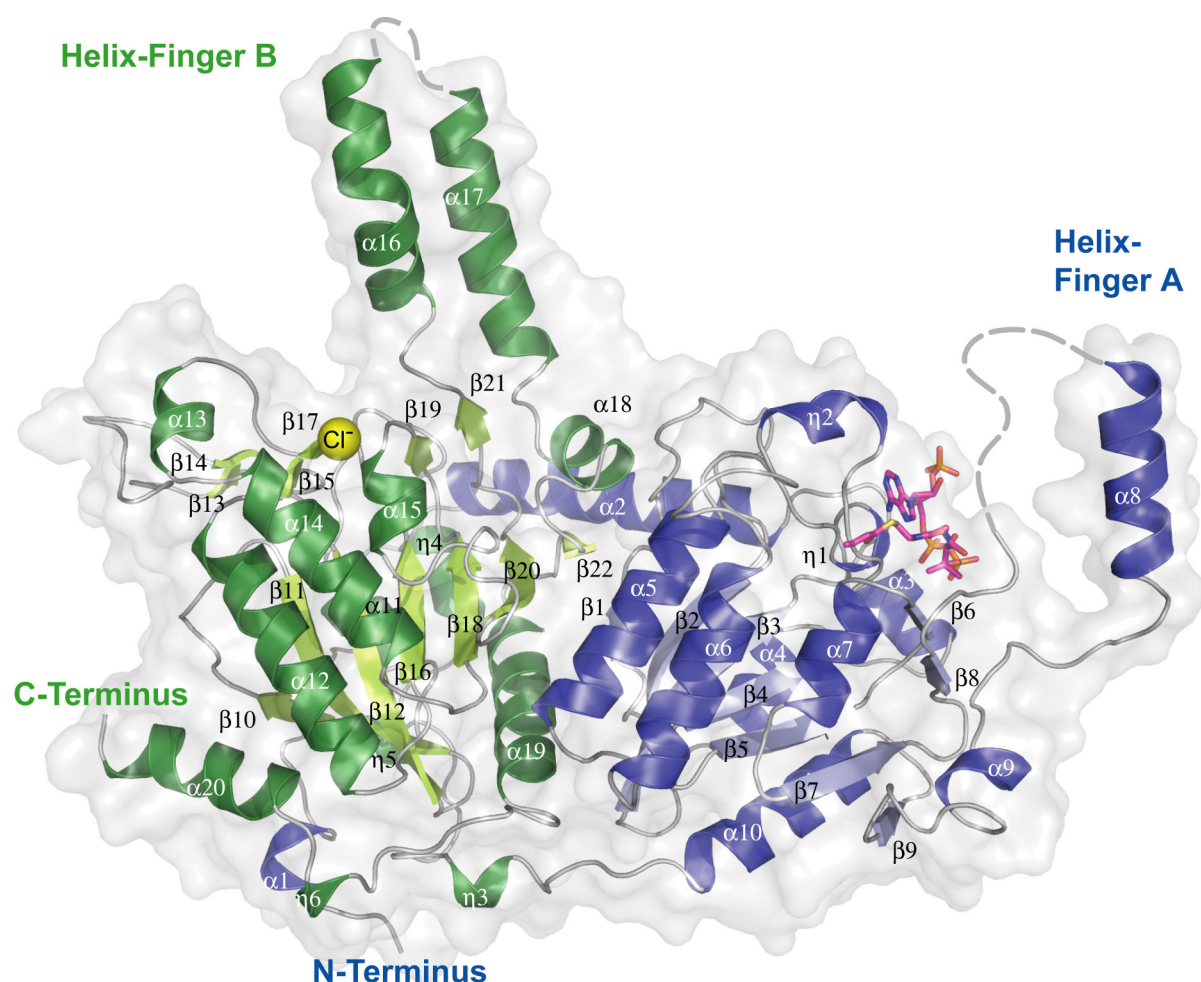
<sup>1</sup> r.m.s.d.  $\triangleq$  root mean square deviation (dt.: Standardabweichung)

<sup>2</sup> Im Falle der Strukturmodelle mit mehr als einem Molekül in der asymmetrischen Einheit sind nur die jeweils größten Lücken angegeben (Nummer der entsprechenden Proteinkette in Klammern).



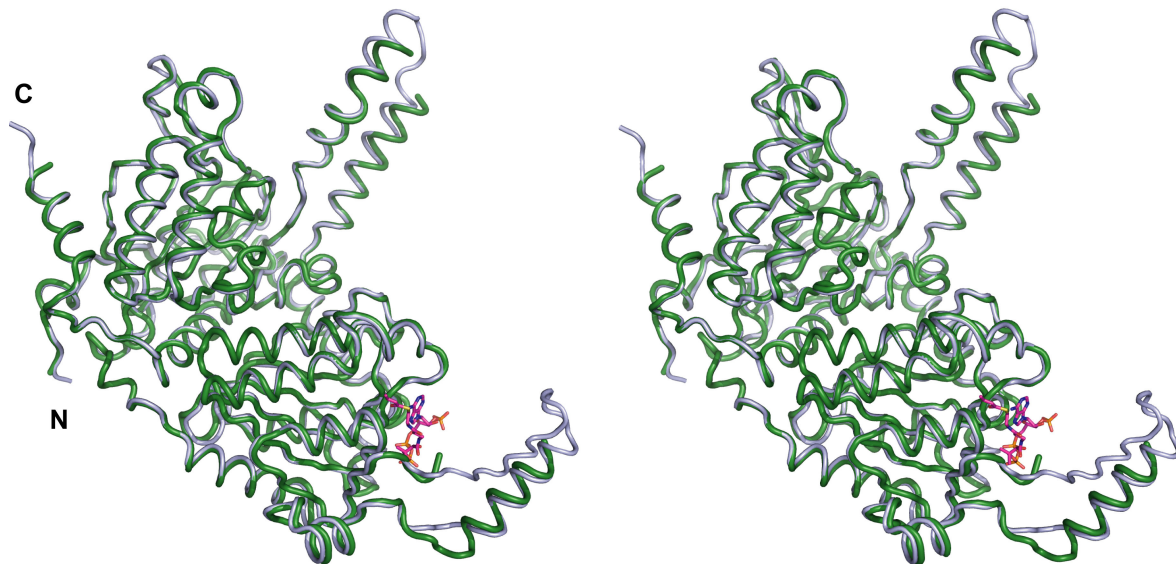
#### 4.3.6 Beschreibung des Strukturmodells für das GcdA-Monomer

Abbildung 4-15 zeigt eine Bänder-Darstellung des GcdA-Monomers mit einem gebundenen Crotonyl-CoA-Molekül. Es hat eine Oberfläche von 25,030 Å<sup>2</sup>. Die Sekundärstrukturmotive wurden mit dem Programm STRIDE [Heinig & Frishman 2004] ermittelt und umfassen insgesamt 20  $\alpha$ -Helices ( $\alpha$ 1- $\alpha$ 20), 6  $3_{10}$ -Helices ( $\eta$ 1- $\eta$ 6) und 22  $\beta$ -Stränge ( $\beta$ 1- $\beta$ 22). Ein mit Hilfe des SSM-Servers [Krissinel & Henrick 2004] durchgeführtes Strukturalignment zeigte die gute Übereinstimmung der *C. symbiosum*-GcdA mit der als Referenzstruktur bei der Strukturlösung verwendeten GcdA aus *A. fermentans*. Die 552 Aminosäurereste, die auf Grundlage



**Abbildung 4-15 Monomer-Struktur der *C. symbiosum*-GcdA** Die Sekundärstrukturmotive wurden mit Hilfe des Programms STRIDE [Heinig & Frishman 2004] zugeordnet. Die  $\alpha$ - und  $3_{10}$ -Helices der N-terminalen Domäne sind in blau und die entsprechenden  $\beta$ -Stränge in hellblau dargestellt. Die Sekundärstrukturmotive der C-terminalen Domäne sind dunkelgrün (Helices) und hellgrün ( $\beta$ -Faltblätter) gefärbt. Das Crotonyl-CoA-Molekül im aktiven Zentrum ist in magenta gezeigt. Aminosäurereste, die nicht durch Elektronendichte definiert waren, sind als gestrichelte graue Linien angedeutet.

ihrer Sekundärstruktur verglichen wurden, wiesen eine Sequenzidentität von 71% auf. Ihre  $C_{\alpha}$ -Atome überlagerten mit einer Standardabweichung von lediglich 0.71 Å. Ein Vergleich der Monomerstrukturen von *A. fermentans*- und *C. symbiosum*-GcdA findet sich in Abbildung 4-16.



**Abbildung 4-16 Stereo-Abbildung einer Überlagerung der Monomer-Strukturmodelle von *C. symbiosum*- (dunkelgrün) und *A. fermentans*-GcdA (hellblau)** Die Farbgebung des dargestellten Crotonyl-CoA-Moleküls entspricht der in Abbildung 4-15.

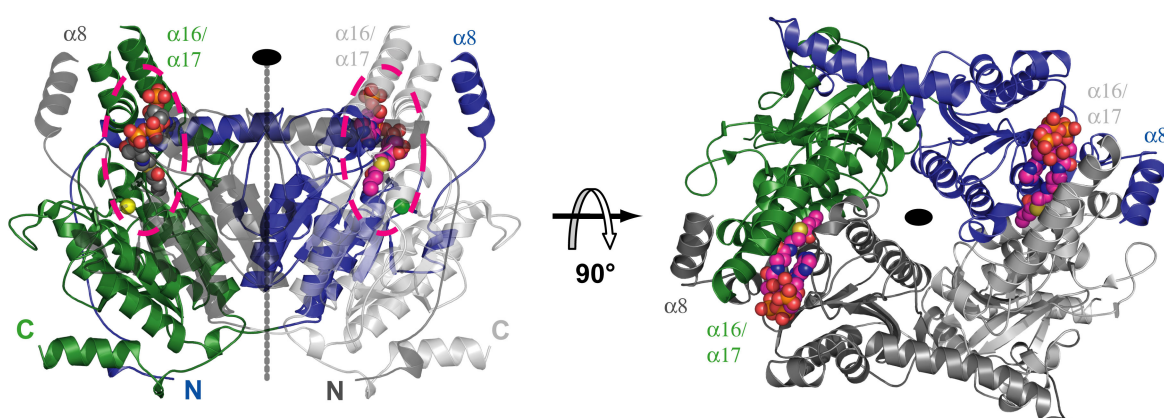
Beide Enzyme gehören zur SCOP-Familie der Biotin-abhängigen Carboxylasen/Carboxyltransferasen, die ihre jeweiligen aktiven Zentren durch zwei verschiedene homologe Unter-einheiten oder Domänen des ClpP/Crotonase-ähnlichen Faltungstyps bilden. Analog zur *A. fermentans*-Struktur [Wendt *et al.* 2003] lässt sich das Monomer der GcdA aus *C. symbiosum* in zwei strukturell und sequenziell äquivalente Domänen von fast gleicher Größe unterteilen: einen N-terminalen Teil, der die Aminosäuren Met3-Ala291 umfasst, und den restlichen C-terminalen Bereich der Polypeptidkette. Die beiden Domänen interagieren über insgesamt 32 Wasserstoffbrückenbindungen und 6 Salzbrücken miteinander. Mehr als die Hälfte dieser Wechselwirkungen sind im N-Terminus der GcdA lokalisiert, wo die Helices  $\alpha 1$  und  $\alpha 2$  die C-terminale Domäne eng umschlungen halten. Das zentrale Strukturmotiv in beiden Domänen ist ein überwiegend paralleles siebensträngiges  $\beta$ -Faltblatt ( $\beta 1$ - $\beta 5$ ,  $\beta 8$  und  $\beta 9$  bzw.  $\beta 10$ - $\beta 12$ ,  $\beta 16$ ,  $\beta 18$ ,  $\beta 20$  und  $\beta 22$ ), welches von sieben benachbarten Helices ( $\alpha 3$ ,  $\alpha 4$ ,  $\alpha 5$ ,  $\alpha 6$ ,  $\alpha 7$ ,  $\alpha 9$



und  $\alpha 10$  bzw.  $\eta 4$ ,  $\alpha 11$ ,  $\alpha 12$ ,  $\alpha 14$ ,  $\alpha 15$ ,  $\alpha 18$  und  $\alpha 20$ ) umschlossen wird. In der C-terminalen Domäne ist dieses zentrale  $\beta$ -Faltblatt orthogonal von sechs kurzen  $\beta$ -Strängen bedeckt, die drei unabhängige doppelsträngige  $\beta$ -Faltblätter ausbilden: zwei parallele ( $\beta 19/\beta 21$  und  $\beta 15/\beta 17$ ) und ein antiparalleles Faltblatt ( $\beta 13/\beta 14$ ). In der N-terminalen Domäne gibt es an der entsprechenden Stelle nur ein paralleles  $\beta$ -Faltblatt ( $\beta 6/\beta 7$ ).

Das auffälligste Strukturmerkmal des GcdA-Monomers sind das  $\alpha$ -Helix-Bündel  $\alpha 16/\alpha 17$ , welches aus der C-terminalen Domäne herausragt, und die  $\alpha$ -Helix  $\alpha 8$ , welche sich in ähnlicher Weise von der N-terminalen Domäne abhebt. Diese Helices werden im Folgenden als Helix-Finger A ( $\alpha 8$ ) und B ( $\alpha 16/17$ ) bezeichnet. Sie sind an der Dimerisierung der GcdA beteiligt und von zentraler Bedeutung für die Architektur des reaktiven Zentrums.

Insgesamt sind die Aminosäurereste aller drei Strukturen gut durch die Elektronendichte definiert. Es gibt allerdings signifikante Ausnahmen, die Helix-Finger betreffend. Eine Reihe von aufeinander folgenden Resten an ihren Enden sind nicht in der Elektronendichte definiert (vgl. Tabelle 4-3, „Lücken“). Die Größe dieser Fehlstellen in der Elektronendichte variiert innerhalb der präsentierten Strukturmodelle und ist in Apo-GcdA am stärksten ausgeprägt. Hier fehlen die Reste Leu502-Gln513 und die komplette Helix  $\alpha 8$  (Gly218-Lys251). Die kleinsten Lücken weist die Glutaconyl-CoA-Kokristallstruktur auf (Gly221-Ala237 und Ala505-Gln513). In Abwesenheit der Liganden sind diese Bereiche – und die Helix  $\alpha 8$  in be-

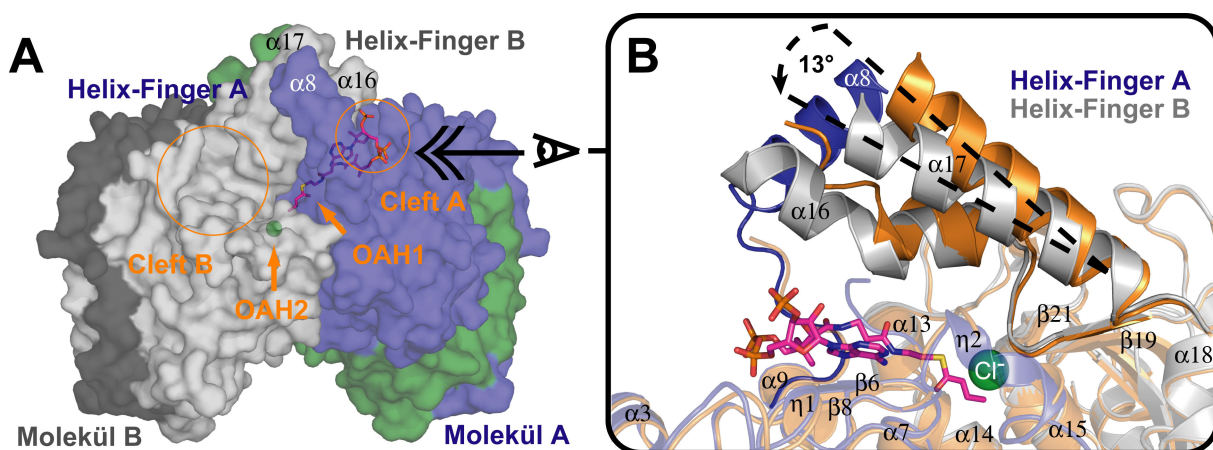


**Abbildung 4-17 Position des aktiven Zentrums im GcdA-Dimer** Die N- und C-terminalen Domänen der symmetrieäquivalenten Monomere sind blau und grün bzw. dunkel- und hellgrau gefärbt. Die Crotonyl-CoA-Moleküle an den intermolekularen Kontaktflächen des Dimers sind als magentafarbenes Chalottenmodell gezeigt. Das zweite Chloridion ist als grüne Kugel dargestellt. Im linken Teil der Abbildung sind die Positionen der aktiven Zentren durch gestrichelte Linien hervorgehoben.

sonderem Maße – strukturell sehr flexibel. Interessanterweise sind die äquivalenten Reste in *A. fermentans*-GcdA gut durch Elektronendichte definiert. Allerdings ist anzunehmen, dass die vergleichsweise hohe strukturelle Integrität der Helices  $\alpha 8$ ,  $\alpha 16$  und  $\alpha 17$  daher rührt, dass sie in diesem Fall direkt an der Ausbildung von Kristallkontakten beteiligt sind, während sie in den hier präsentierten *C. symbiosum*-GcdA-Strukturen frei beweglich in den Lösungsmittelraum ragen.

#### 4.3.7 Architektur des aktiven Zentrums der GcdA

Das aktive Zentrum der GcdA setzt sich durch eine C2-symmetrische „Kopf-zu-Schwanz“-Dimerisierung aus den Molekülen A und B zusammen, wie in Abbildung 4-17 gezeigt. Dabei kommt jeweils die N-terminale Domäne des einen Monomers auf der C-terminalen Domäne des zweiten Moleküls zum Liegen. An den Kontaktflächen der Domänen entstehen auf diese Weise die enzymatischen Taschen. Sie werden von den ineinander verschränkten Helix-Fingern A und B wie von einem Deckel abgeschlossen (Abbildung 4-18). Diese Bi-Domänen-Architektur der Reaktionszentren ist weit verbreitet innerhalb der Crotonase-Superfamilie [Hamed *et al.* 2008] und charakteristisch für alle strukturell charakterisierten Biotin-abhängigen Decarboxylasen/Carboxyltransferasen [Diacovich *et al.* 2004].



**Abbildung 4-18 Architektur des aktiven Zentrums** (A) Oberflächendarstellung des GcdA-Dimers mit Blick auf die „verschränkten“ Helix-Finger. Der Farbcode ist identisch mit dem in Abbildung 4-17. Die orangenen Kreise zeigen die Position der beiden Vertiefungen an, die zu OAH1 und OAH2 führen. (B) Überlagerung des Crotonyl-CoA-GcdA-Komplexes mit dem Apo-GcdA-Strukturmodell. Dargestellt ist eine Vergrößerung der Glutaconyl-CoA-Bindungsstelle an der Kontaktfläche des Dimers. Die entsprechenden Polypeptidketten des Apo-Modells sind orange gefärbt. Die gestrichelten Linien verdeutlichen die „Bewegung“, die Helix  $\alpha 17$  beim Binden des Substrats beschreibt.

Zu beiden Seiten der „verschränkten Finger“ befinden sich Vertiefungen auf der Proteinoberfläche. Diese werden im Folgenden als Cleft A und Cleft B bezeichnet. Sie bilden die Zugänge zu zwei verschiedenen Oxyanionen-Löchern (engl.: *oxyanion holes*, OAH) am Boden der aktiven Zentren. Cleft A hat die Form einer tiefen Spalte und führt zum N-terminalen Teil der enzymatischen Tasche, der das Oxyanionen-Loch OAH1 (Val153, Gly196) und die Glutaconyl-CoA-Bindungsstelle enthält. In den Substratkomplex-Strukturen konnte die zusätzliche Elektronendichte in diesem Teil des aktiven Zentrums eindeutig dem Substratanalogon Glutaconyl-CoA und dem Produkt der katalysierten Reaktion, Crotonyl-CoA, zugeordnet werden (vgl. Abbildung 4-14).

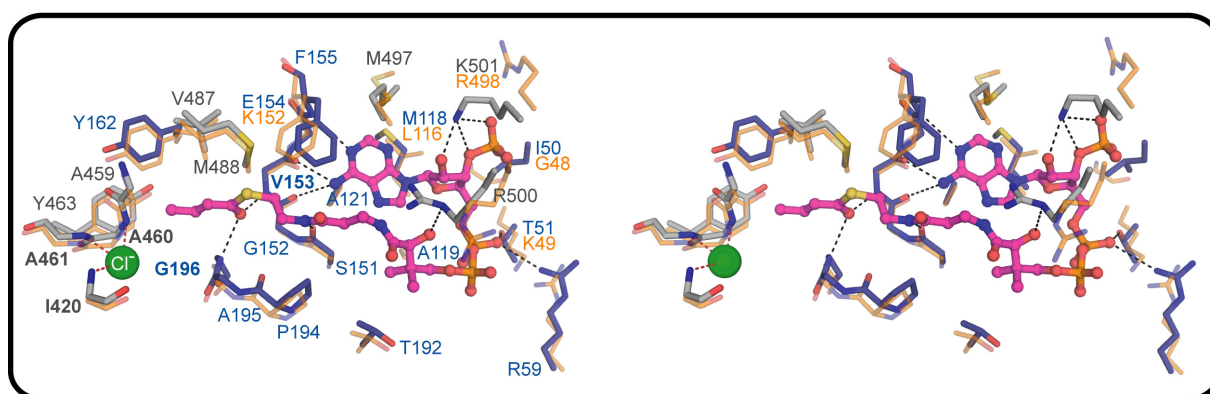
Die trichterförmige Vertiefung Cleft A stellt den Zugang zur wahrscheinlichen Bindungstasche der prosthetischen Gruppe Biotin dar und enthält das OAH2 (Ile420, Ala460, Ala461). Diese Kavität ist deutlicher ausgeprägt als Cleft A und vermutlich die Kontaktfläche für die Interaktion mit dem Biotin-Carrier GcdC [Wendt *et al.* 2003]. Leider ist es im Rahmen dieser Arbeit trotz umfangreicher Kokristallisations- und *soaking*-Experimente nicht gelungen, eine Biotin- bzw. Biocytin-GcdA-Komplexstruktur zu erhalten. Allerdings wurde beobachtet, dass in Apo-GcdA und in der Glutaconyl-CoA-Kokristallstruktur OAH2 von einem Wassermolekül besetzt ist, während im Falle der Glutaconyl-CoA- bzw. Crotonyl-CoA-Komplexe die zusätzliche Elektronendichte an dieser Stelle einem Chloridion zugeordnet werden konnte.

Die Überlagerung der Crotonyl-CoA-Komplexstruktur mit Apo-GcdA offenbart einige bemerkenswerte strukturelle Veränderungen, die offenbar aus der Bindung des Substrates in der enzymatischen Tasche resultieren (Abbildung 4-18B). Die auffälligste dieser Modifikationen ist das Auftauchen des Helix-Fingers A ( $\alpha 8$ ), der in der Elektronendichte der Apo-GcdA-Struktur nur undeutlich zu erkennen war. Darüber hinaus bewegt sich  $\alpha 17$  im Vergleich zur Apo-Struktur um ca.  $13^\circ$  in Richtung des gebundenen Crotonyl-CoA. Diese Bewegung ist möglicherweise Teil eines dynamischen Prozesses, der die Transcarboxylierungsreaktion mit konformationellen Änderungen an der Oberfläche des GcdA-Subkomplexes koppelt.

#### 4.3.8 Die Glutaconyl-CoA-Bindungstasche

Der Boden der Glutaconyl-CoA-Bindungstasche (Cleft A) wird durch die  $3_{10}$ -Helices  $\eta 1$  und  $\eta 2$ , die diese verbindenden Schleifenregionen und den  $\beta$ -Strang  $\beta 6$  gebildet. Abgeschlossen wird die Tasche durch die gekreuzten Helix-Finger, die sich auf das Substrat legen wie ein

Deckel. Eine Reihe von positiv geladenen Aminosäureresten der  $\alpha$ -Helices  $\alpha 3$  und  $\alpha 16$  „bewachen“ den Eingang zur ansonsten größtenteils ungeladenen Bindungstasche. Die Substrat-Protein-Wechselwirkungen sind dementsprechend überwiegend hydrophober Natur und gehen hauptsächlich von Resten der N-terminalen Domäne aus. Der detaillierte Bindungsmodus der CoA-Ester in *C. symbiosum*-GcdA ist anhand von Crotonyl-CoA in Abbildung 4-19 gezeigt.



**Abbildung 4-19 Stereo-Abbildung der Glutaconyl-CoA-Bindestelle** Dargestellt sind die Aminosäurereste im Abstand von 4 Å zum gezeigten Crotonyl-CoA-Molekül. Reste der N- und C-terminalen Domäne sind blau bzw. grau eingefärbt. Strukturell äquivalente Aminosäuren der *A. fermentans*-GcdA sind in orange abgebildet, nicht-konservierte Reste in orange beschriftet. Die gestrichelten Linien deuten Wasserstoffbrückenbindungen an.

Die U-förmige Konformation des CoA-Anteils ähnelt stark der Konformation anderer Biotin-abhängiger Carboxylasen, wie z.B. der Carboxyltransferase-Domäne der Acetyl-CoA-Carboxylase [Zhang *et al.* 2003] und der Transcarboxylase 12S [Hall *et al.* 2003]. Im Vergleich zur *A. fermentans*-Struktur sind die meisten der spezifischen Interaktionen konserviert (vgl. Abbildung 4-19). Der Adenin-Teil von CoA steckt in einer hydrophoben Mulde, die von Met118 und Ala121 einerseits und Phe155 und Met497 andererseits gebildet wird. Die Pantothenat-Einheit ist tief vergraben in einer Tasche, die aus Aminosäureresten von  $\beta 6$ ,  $\eta 2$  und  $\alpha 16$  besteht. Sowohl das Ribose-3'-Phosphat als auch das Pyrophosphat sind stärker Lösungsmittel-exponiert und daher strukturell flexibler, was sich im Vergleich zum Rest des Moleküls in deutlich erhöhten B-Faktoren bemerkbar macht. In diesem Bereich gibt es eine Reihe von Resten, die über Wasserstoffbrückenbindungen mit den Phosphatgruppen des CoA-Esters interagieren (u.a. Lys501 und Arg59). Der Acetyl-Anteil des CoA-Esters steckt tief in der Bindungstasche, wobei der Carbonyl-Sauerstoff fest an OAH1 bindet. In der Glu-

taryl-CoA-Kokristallstruktur zeigt die Carboxylgruppe in Richtung des OAH<sub>2</sub>, welches wahrscheinlich den Ureido-Sauerstoff des Kosubstrats Biotin bindet.

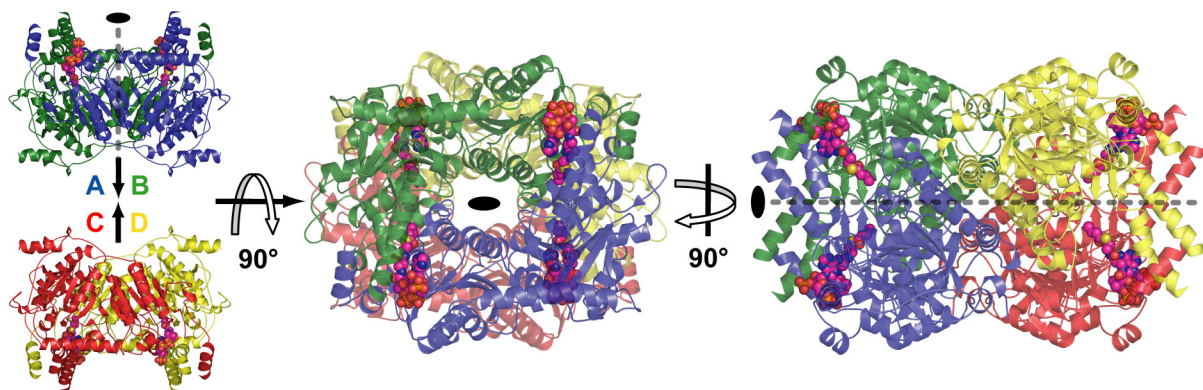
### 4.3.9 Quartärstruktur

Der *A. fermentans*-Struktur zufolge bildet GcdA C<sub>2</sub>-symmetrische Dimere des in Abbildung 4-17 gezeigten Aufbaus [Wendt *et al.* 2003]. Allerdings eluiert die GcdA aus *C. symbiosum* bei der Aufreinigung mittels Volumenausschlusschromatographie bei einem apparenten Molekulargewicht von 232 kDa, was auf eine tetramere Zusammensetzung hinweist (berechnetes Molekulargewicht: 260 kDa). Darüber hinaus zeigte eine Analyse der Kristallpackung der *C. symbiosum*- und der *A. fermentans*-Strukturen, dass sie alle aus identischen 222-symmetrischen Tetrameren aufgebaut sind. Diese Befunde wurden zusätzlich durch die Ergebnisse folgender webbasierter Programme bestätigt, welche die Quartärstruktur von Proteinen voraussagen: *Protein Interfaces and Assemblies-Server* [Ponstingl *et al.* 2003], *Protein Quaternary Structure-Server* und *Protein Interfaces, Surfaces and Assemblies Service* [Krissinel & Henrick 2007]. Auf Grundlage der jeweiligen Strukturkoordinaten sagten alle diese Programme unabhängig voneinander voraus, dass *C. symbiosum*- und *A. fermentans*-GcdA stabile Tetramere bilden.

Die Enzyme der Crotonase-Superfamilie nehmen eine Reihe unterschiedlicher Quartärstrukturen an [Hamed *et al.* 2008], wie z.B. Monomere (1fc7: Photosystem II D1 CTPase [Liao *et al.* 2000]), Dimere (1od4: Carboxyltransferase-Untereinheit der Acetyl-CoA Carboxylase aus Hefe [Zhang *et al.* 2003]), Hexamere (1on3: Transcarboxylase 12S [Hall *et al.* 2003]) und sogar Tetradecamere (1tyf: proteolytische Untereinheit der caseinolytischen Protease [Wang *et al.* 1997]). Interessanterweise sind die tetrameren Carboxyltransferasen der Acetyl-CoA-Carboxylase aus *Staphylococcus aureus* und *Escherichia coli* (2f9i und 2f9y [Bilder *et al.* 2006]), die ebenfalls zur strukturellen Unterfamilie der Biotin-abhängigen Carboxylasen gehören, in ihrer Quartärstruktur nicht mit GcdA verwandt.

Das 222-symmetrische GcdA-Tetramer erhält man durch einfache Rotation um die kristallographischen Symmetrieachsen. Es hat eine Oberfläche von 68,280 Å<sup>2</sup>, wobei insgesamt 36,310 Å<sup>2</sup> Lösungsmittel-zugängliche Oberfläche (9,078 Å<sup>2</sup> pro Monomer) verdeckt werden. Das Tetramer entsteht, wie in Abbildung 4-20 gezeigt, durch die Assoziation der identischen Dimere AB und CD. Dabei entstehen jeweils Berührungsflächen von 4,232 Å<sup>2</sup>. Die Kombination dieser Dimere generiert weitere Kontaktflächen von 1,743 Å<sup>2</sup> und 1,307 Å<sup>2</sup> entlang der AC/BD- bzw. AD/BC-Berührungsflächen. Das Tetramer wird durch insgesamt 104 Wasser-

stoffbrückenbindungen stabilisiert. Mehr als 95% dieser Wechselwirkungen sind im Vergleich zur *A. fermentans*-Struktur konserviert. Jede der AB- und CD-Kontaktflächen enthält 38 Wasserstoffbrückenbindungen und zwei Salzbrücken zwischen His261 und Glu427. Ein beträchtlicher Teil der Wasserstoffbrücken geht von den Helices  $\eta 2$  und  $\alpha 6$  und der diese Helices verbindenden Schleifenregion aus. Diese interagieren mit  $\alpha 5$  und  $\alpha 18$  sowie vor allem mit  $\alpha 15$  und den angrenzenden Schleifenregionen des anderen Monomers. Bemerkenswert ist, dass es zwischen den Helix-Fingern  $\alpha 8$  und  $\alpha 16$  nur eine Wasserstoffbrückenbindung gibt, obwohl sie durch hydrophobe Wechselwirkungen stark zur Bildung der AB- bzw. CD-Dimere beitragen. Die AC/BD-Grenzflächen beinhalten 30 Wasserstoffbrücken, die durch die Interaktion des Strukturmotivs  $\alpha 9/\beta 9/\alpha 10$  mit dem N-Terminus und  $\eta 3$  sowie durch die Wechselwirkung von  $\alpha 6$  und  $\alpha 7$  mit  $\eta 6$  zu Stande kommen. Die übrigen Dimer-Kontaktflächen AD und BC beschränken sich auf die C-terminalen Domänen der Monomere und enthalten 36 Wasserstoffbrücken und Salzbrücken zwischen Arg578 und Glu430 sowie Arg391 und Glu582. Die wichtigste Wechselwirkung dieser Grenzflächen besteht zwischen dem Helix-Bündel  $\alpha 12/\alpha 14$  und  $\alpha 20$ .



**Abbildung 4-20** Zwei orthogonale Orientierungen des 222-symmetrischen GcdA-Tetramers Die Ketten A bis D sind blau, grün, rot bzw. gelb eingefärbt.



## 5 Diskussion

### 5.1 Enamidase

#### 5.1.1 Vergleich der Enamidase mit anderen Amidohydrolasen

Tabelle 5-1 zeigt das Ergebnis eines Sekundärstrukturvergleichs des Enamidase-Monomers (Kette A) mit ähnlichen Strukturmodellen aus der RCSB PDB. Das Alignment wurde mit dem SSM-Server durchgeführt (<http://www.ebi.ac.uk/msd-srv/ssm/cgi-bin/ssmserver>) [Krissinel & Henrick 2004], wobei zunächst identische Treffer aussortiert und nur vollständig charakterisierte Strukturen miteinbezogen wurden, um die Signifikanz des Vergleichs zu erhöhen. Alle in der Liste enthaltenen Enzyme gehören der Amidohydrolase-Familie an. Wie erwartet, zeigen die Proteine eine relativ hohe strukturelle Homologie (die Standardabweichung der Position ihrer C $\alpha$ -Atome liegt zwischen 2.3 und 3.2 Å), während sie nur eine recht geringe Sequenzähnlichkeit aufweisen (11-21% Sequenzidentität unter den alignierten Resten). Ein strukturbasiertes Sequenzalignment der Enamidase mit ihren engsten Strukturverwandten *Aquifex aeolicus*-Dihydroorotase (Aa.DHO), *E. coli*-Isoaspartyl-Dipeptidase (Ec.IAD) und *Alcaligenes faecalis*-D-Aminoacylase (Af.DAA), das ebenfalls mit Hilfe des SSM-Servers erstellt wurde, unterstreicht diesen Befund (Abbildung 5-1): Abgesehen von den Histidin- und Aspartat-Resten des Metallzentrums gibt es nur wenige Gemeinsamkeiten in der Sequenz der vier Enzyme.

Nach der von Seibert und Raushel eingeführten Klassifikation der Amidohydrolasen (vgl. 1.1.3) stellt Enamidase einen Spezialfall der Typ II-Amidohydrolasen dar, da sie der erste Vertreter dieser Klasse ist, der ein binukleares Metallzentrum hat, das sowohl Zink als auch Eisen enthält [Seibert & Raushel 2005]. Neben dem Phosphotriesterase-ähnlichen Protein aus *E. coli* ist Enamidase die einzige Typ II-Amidohydrolase, für die eine Kristallstruktur verfügbar ist. Interessanterweise gehören alle anderen Amidohydrolasen, die zyklische Substrate hydrolysieren, wie beispielsweise die *E. coli*-Dihydroorotase und alle strukturell charakterisierten Dihydropyrimidinasen (Hydantionasen), zum häufigsten Typ I, in dem das Metallverbrückende Glutamat durch ein carboxyliertes Lysin ersetzt ist.

Protein <sup>a</sup>	Organismus	PDB	rmsd (Å)	Nalign <sup>b</sup>	%seq <sup>c</sup>	Metall	Klasse <sup>d</sup>
<b>Enamidase</b> [Kress <i>et al.</i> 2008]	<i>E. barkeri</i>	2vun:A	-	385	100	$\alpha\beta$ FeZn	II*
<b>Dihydroorotase</b> [Martin <i>et al.</i> 2005]	<i>A. aeolicus</i>	1xrt:A	2.33	293	21	$\alpha$ Zn	n/a
<b>Isoaspartyl-Dipeptidase</b> [Jozic <i>et al.</i> 2003]	<i>E. coli</i>	1poj:B	2.35	306	18	$\alpha\beta$ Zn	I
<b>D-Aminoacylase</b> [Liaw <i>et al.</i> 2003]	<i>A. faecalis</i>	1m7j:A	2.41	288	21	$\beta$ Zn	V
<b>D-Hydantionase</b> [Abendroth <i>et al.</i> 2002]	<i>Thermus</i> sp.	1gkp:A	2.45	298	21	$\alpha\beta$ Zn	I
<b>NAG-6-P-Deacetylase</b> [Hall <i>et al.</i> 2007]	<i>E. coli</i>	2p50:C	2.50	291	16	$\beta$ Zn	n/a
<b>Phosphotriesterase-ähnliches Protein</b> [Buchbinder <i>et al.</i> 1998]	<i>E. coli</i>	1bf6:B	2.52	197	16	$\alpha\beta$ Zn	II
<b>L-Hydantionase</b> [Abendroth <i>et al.</i> 2002]	<i>A. aurescens</i>	1gkr:A	2.55	300	20	$\alpha\beta$ Zn	I
<b>Dihydropyrimidinase</b> [Lohkamp <i>et al.</i> 2006]	<i>S. kluyveri</i>	2fty:A	2.75	297	20	$\alpha\beta$ Zn	I
<b>Dihydropyrimidinase</b> [Lohkamp <i>et al.</i> 2006]	<i>D. discoideum</i>	2ftw:A	2.75	301	16	$\alpha\beta$ Zn	I
<b>CRMP1</b> [Deo <i>et al.</i> 2004]	<i>Mus musculus</i>	1kcx:B	2.75	306	18	none	n/a
<b>Cytosin-Deaminase</b> [Ireton <i>et al.</i> 2002]	<i>E. coli</i>	1k70:A	2.80	273	16	$\alpha$ Fe	III
<b>Dihydroorotase</b> [Thoden <i>et al.</i> 2003]	<i>E. coli</i>	1j79:A	2.84	239	20	$\alpha\beta$ Zn	I
<b>NAG-6-P-Deacetylase</b> [Vincent <i>et al.</i> 2004]	<i>B. subtilis</i>	2vhl:A	2.86	291	17	$\alpha\beta$ Fe	IV
<b>SIH-Deaminase</b> [Hermann <i>et al.</i> 2007]	<i>T. maritima</i>	2plm:A	2.95	293	14	$\alpha$ Zn	n/a
<b>Phosphotriesterase</b> [Benning <i>et al.</i> 1995]	<i>P. diminuta</i>	1psc:B	3.17	181	11	$\alpha\beta$ Zn	I

Tabelle 5-1 Sekundärstrukturalignment der Enamidase

<sup>a</sup> Das Strukturalignment wurde mit dem SSM-Server (<http://www.ebi.ac.uk/msd-srv/ssm/cgi-bin/ssmserver>) durchgeführt. Dabei wurden die Standardeinstellungen verwendet und die 24 besten Treffer anschließend zur Verbesserung der Signifikanz manuell sortiert, um die Redundanz zu verringern.

<sup>b</sup> Länge des Alignments Nalign. Der dreidimensionale Vergleich der Referenz- und Zielstrukturen wird auf der Grundlage ihrer räumlichen Nähe vorgenommen, wobei die Standardabweichung der C $\alpha$ -Atome minimiert und die Anzahl der alignierten Reste maximiert wird.

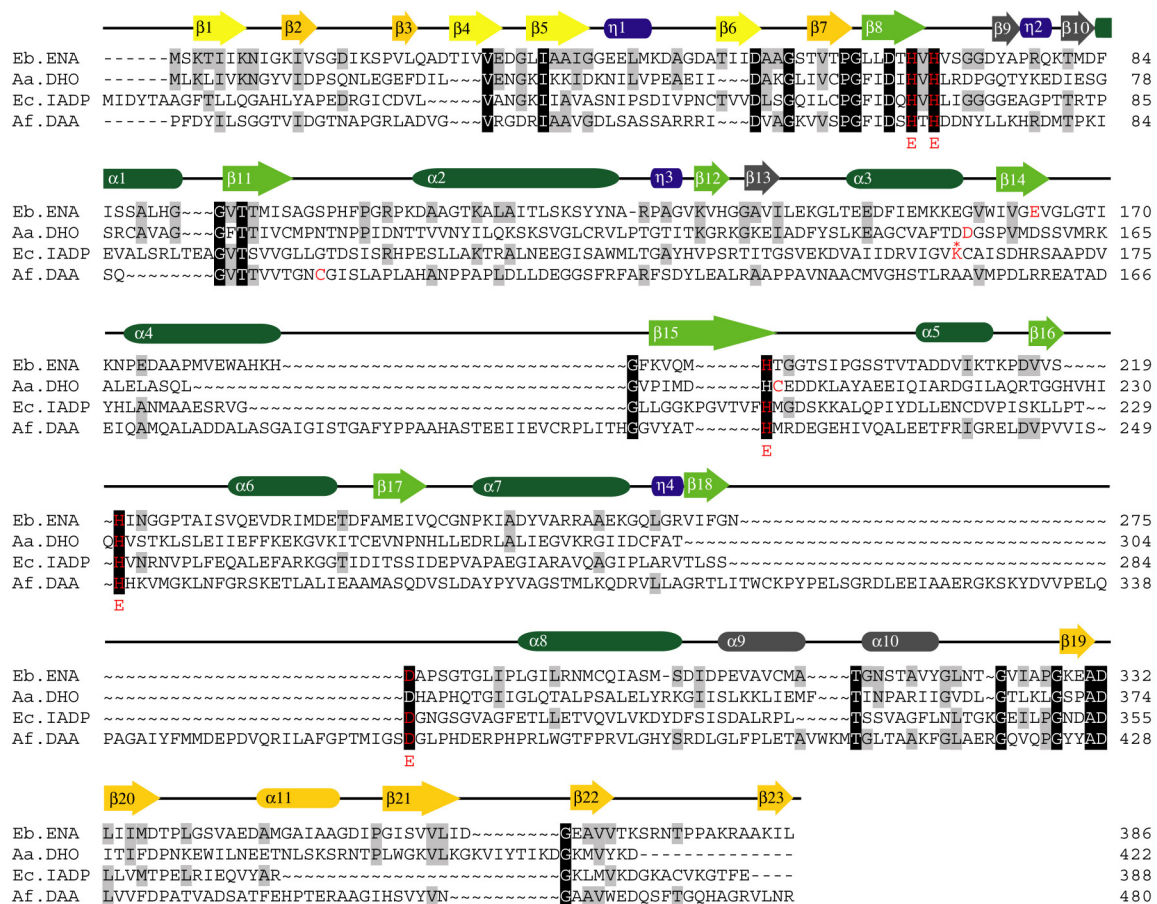
<sup>c</sup> Sequenzidentität %seq ist der prozentuale Anteil von Paaren identischer Reste Nident an der Gesamtzahl alignierter Aminosäuren Nalign: %seq = Nident / Nalign.

<sup>d</sup> Die Amidohydrolasen-Familie kann anhand der Zusammensetzung des reaktiven Zentrums in sieben verschiedene Typen unterteilt werden (siehe Text) [Seibert & Raushel 2005].

\* Das Metallzentrum von Enamidase enthält sowohl Fe<sup>2+</sup> als auch Zn<sup>2+</sup> und stellt damit einen Spezialfall einer Typ II-Amidohydrolase dar.



Abbildung 5-2 zeigt eine Überlagerung der aktiven Zentren von Enamidase und den strukturverwandten Enzymen Aa.DHO, Ec.IAD und Af.DAA. Bemerkenswert ist, dass die räumliche Anordnung der Metallliganden außergewöhnlich gut übereinstimmt, obwohl die jeweiligen Proteinketten auf der Grundlage ihrer Sekundärstruktur superpositioniert wurden. Alle vier gezeigten Metallzentren gehören jeweils verschiedenen Typen an (vgl. Tabelle 5-1). Entsprechend findet sich nur das stark konservierte HXH-Motiv in allen vier Enzymen wieder, während die Identität des Metall-verbrückenden Aminosäurerestes in allen vier Strukturen unterschiedlich ist.



**Abbildung 5-1 Strukturbasiertes Sequenzalignment der Enamidase und verwandter Enzyme**

Das Alignment wurde mit Hilfe des SSM-Server durchgeführt. Die Sekundärstrukturelemente der Enamidase sind als Pfeile (β-Stränge) und abgerundete Kästchen (α- und 310-Helices) dargestellt. Sie sind entsprechend mit β, α und η beschriftet und jeweils fortlaufend nummeriert. Der Farbcode entspricht Abbildung 4-5. Aminosäuren, die in allen vier Sequenzen konserviert sind, sind schwarz, Reste, die identisch mit der Enamidase-Sequenz sind, grau hervorgehoben. Rote Buchstaben zeigen die Reste an, die jeweils in die Metallbindung involviert sind. Großgeschriebene rote E's unterhalb des Sequenzalignments zeigen die Position der Metallbindenden Reste der Enamidase an.

Ec.IAD gehört zum archetypischen Typ I und beinhaltet ein carboxyliertes Lysin, das die Zinkionen seines binuklearen Metallzentrums verbrückt. Af.DAA (Typ V) enthält stattdessen ein Cystein an dieser Stelle. Aa.DHO hingegen stellt einen völlig neuen Typus dar: Hier wird die Stelle des Metall-verbrückenden Restes von einer Aspartat-Seitenkette eingenommen und die  $\beta$ -Position des Metallzentrums von einem Cystein-Rest besetzt. Obwohl das strukturbasierte Sequenzalignment (vgl. Abbildung 5-1) den Resten His193, His220 und Asp276 der Enamidase strukturell äquivalente Reste der Aa.DHO-Sequenz zuordnet, befinden sich die entsprechenden Reste zwar in der Nähe des mononuklearen Metallzentrums der Aa.DHO, partizipieren jedoch nicht an der Metallbindung.

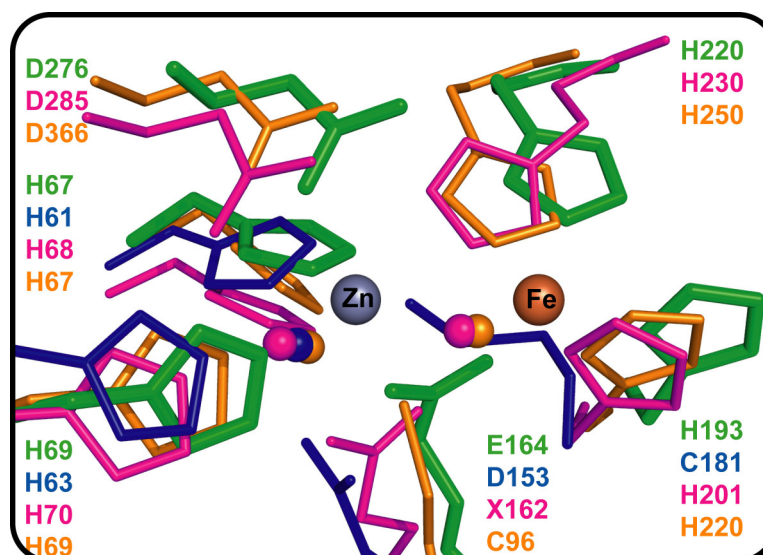
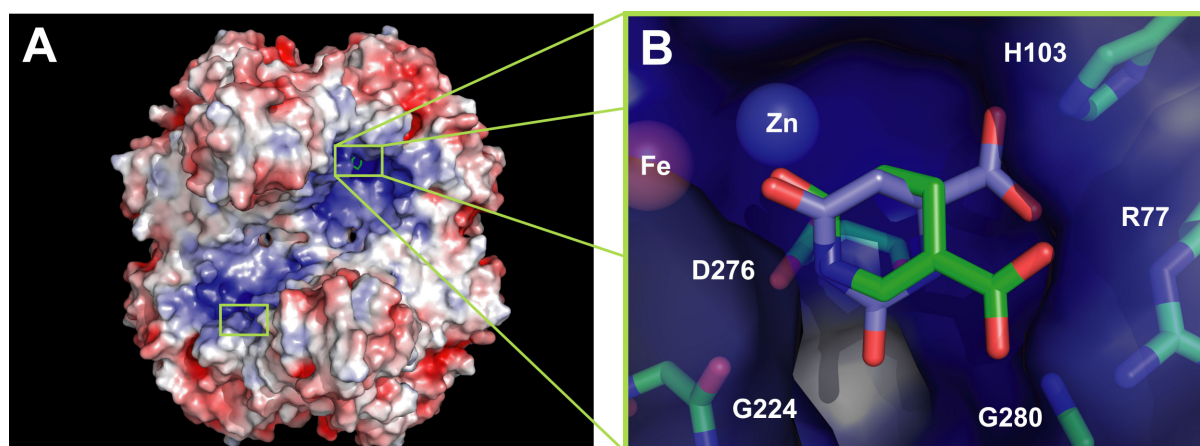


Abbildung 5-2 Strukturbasiertes Alignment des aktiven Zentrums von Enamidase und strukturell verwandten Enzymen Aa.DHO (1xrt), blau; Ec.IAD (1poj), magenta; Af.DAA (1m7j), orange. Die Zink- und Eisenionen der Enamidase sind in blau und orange dargestellt.

### 5.1.2 Substraterkennung

Wie die graphische Darstellung des elektrostatischen Oberflächenpotentials der Enamidase in Abbildung 5-3A zeigt, ist die Oberfläche der Enamidase größtenteils negativ geladen. Dies steht im Einklang mit dem aus der Aminosäuresequenz berechneten pI von 5.2 und einer vorausgesagten Nettoladung von -10.7 bei pH 7.0. Die Ausnahme bildet eine trichterförmige Vertiefung an der Spitze des ( $\alpha/\beta$ )-Fasses, die zum aktiven Zentrum an ihrem Boden führt.

Auf Grund der gebundenen Metallionen und der positiv geladenen Reste in ihrer direkten Nachbarschaft wird dieser Bereich von positiver Oberflächenladung dominiert und ist damit komplementär zum anionischen Substrat THON.



**Abbildung 5-3 Substraterkennung** (A) Elektrostatisches Oberflächenpotential des Enamidase-Tetramers, berechnet mit dem Programm APBS (rot:  $-5 \text{ kcal} \cdot \text{e}^{-1}$ , weiß:  $0 \text{ kcal} \cdot \text{e}^{-1}$ , blau:  $+5 \text{ kcal} \cdot \text{e}^{-1}$ ,  $0.1 \text{ M}$  Salz) [Baker *et al.* 2001]. Die Positionen von zwei der aktiven Zentren sind durch grüne Kästchen hervorgehoben. Die Orientierung entspricht der linken Seite von Abbildung 4-6. (B) Bindungsmodell für THON an das binukleare Metallzentrum der *E. barkeri*-Enamidase. THON und Dihydroorotat sind in grün bzw. hellblau dargestellt.

Die Versuche, durch Kokristallisations- und *soaking*-Experimente eine Komplexstruktur von Enamidase mit THON zu erhalten, sind leider erfolglos geblieben. Obwohl das Enzym in Gegenwart hoher THON-Konzentrationen problemlos Kristalle von guter Qualität bildete, lieferte die Auswertung zahlreicher Datensätze keine signifikante Elektronendichte für Substrat oder Produkt. Es wird vermutet, dass dies hauptsächlich an der Hydrolyse-Empfindlichkeit des Substrates THON bzw. der spontanen Decarboxylierung des Produktes 2-Formylglutarat liegt [Bozarth 2005]. Um dennoch Einblick in die Substraterkennung und den Mechanismus der Enamidase zu erhalten, wurde nach passenden Modellen unter den verwandten Enzymen der Amidohydrolase-Familie gesucht. Auf Grund der Verfügbarkeit eines hochaufgelösten Modells eines Edukt- und Produkt-Komplexes und der strukturellen Ähnlichkeit ihrer jeweiligen Substrate wurde die *E. coli*-Dihydroorotase (Ec.DHO; PDB-Code: 1j79, [Thoden *et al.* 2003]) als Modell für die Bindung von THON durch Enamidase ausgewählt.

Nachdem die Strukturen von Ec.DHO und Enamidase durch Vergleich ihrer Sekundärstruk-

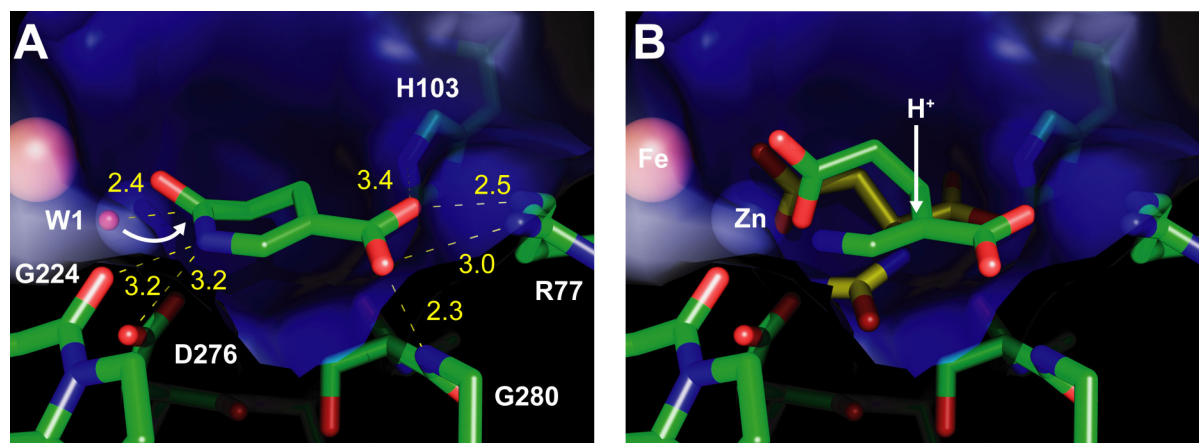
turen überlagert worden waren, wurde das gebundene Dihydroorotat-Molekül als Ausgangspunkt für die manuelle Positionierung von THON verwendet (Abbildung 5-3B). Mit Hilfe dieses Modells für den Enamidase-THON-Komplex konnte eine Reihe von Resten im aktiven Zentrum identifiziert werden, die möglicherweise eine Rolle bei der Erkennung des Substrates durch das Enzym spielen. Im finalen Modell koordiniert der Carbonyl-Sauerstoff von THON an das Zinkion (2.7 Å), während das zugehörige Kohlenstoffatom ca. 2.4 Å von dem Metall-verbrückenden Wassermolekül W1 (Abbildung 5-4A) entfernt ist. Die Carboxylat-Gruppe von THON scheint durch elektrostatische Wechselwirkung mit den Seitenketten der Aminosäuren Arg77 und His103 sowie mit dem Stickstoffatom der Peptidbindung von Gly280 in Position gehalten zu werden. Die Abstände liegen dabei zwischen 2.5 (Gly280) und 3.4 Å (His103). Mit ca. 3.2 Å ist der Amid-Stickstoff von THON in Wasserstoffbrückenbindungsabstand zu einem der Carboxylat-Sauerstoffatome der Asp276-Seitenkette sowie zum Sauerstoff der Peptidbindung von Gly224. Die letztgenannte Wechselwirkung ist möglicherweise entscheidend für die Substraterkennung, da sie durch die Fixierung der Orientierung von THON im aktiven Zentrum die Stereospezifität der hydrolytischen Reaktion gewährleistet.

### 5.1.3 Der Mechanismus der Enamidase-Reaktion

Basierend auf diesem Bindungsmodell und den Ähnlichkeiten zwischen den von Ec.DHO und Enamidase katalysierten Reaktionen wird folgender Mechanismus vorgeschlagen: Im ersten Schritt der Reaktion deprotoniert Asp276 das verbrückende Wassermolekül W1, während das exponiertere Eisenion die Carbonyl-Sauerstoffbindung polarisiert und damit den nukleophilen Angriff von W1 auf das Kohlenstoffatom des Amids erleichtert. Das entstehende tetraedrische Intermediat wird durch die zweizählige Koordination zu beiden Ionen des Metallzentrums stabilisiert. Begünstigt durch die zeitgleiche Protonierung des Amid-Stickstoffatoms durch Asp276 wird die Kohlenstoff-Stickstoffbindung gespalten, womit dieser Teil der Zweistufenreaktion abgeschlossen wird.

Analog zum Carbamoylaspartat in der Ec.DHO-Struktur verbrückt das neu gebildete Carboxylat des Reaktionsintermediats 2-(Enamin)-glutarat die Zink- und Eisenionen des Metallzentrums (Abbildung 5-4B). Da dieses Enamin anfällig ist für die spontane Hydrolyse durch umgebende Wassermoleküle, stellt die Enantioselektivität der folgenden Reaktion die eigentliche Leistung der Enamidase dar. Wie in Abbildung 1-3 illustriert, geht dieser zweiten Hyd-

rolyse die elektrophile Addition eines Protons von der *Re*-Seite des 2-(Enamin)-glutarats voraus, wobei (*S*)-2-(Formimino)-glutarat (Abbildung 5-4B) entsteht. Die resultierende positive Ladung wird stabilisiert durch elektrostatische Interaktion mit dem Sauerstoff der Peptidbindung von Gly224. Die Addition eines Wassermoleküls, gefolgt von der Eliminierung von Ammonium, führt schließlich zur Bildung von (*S*)-2-Formylglutarat.



**Abbildung 5-4 Mechanismus der von Enamidase katalysierten Zweistufenreaktion** (A) Modell der Substraterkennung. Das putative Nukleophil W1 ist als rote Kugel dargestellt. Die interatomaren Abstände (gelbe Ziffern) sind in Å angegeben. (B) Situation nach der Ring-öffnenden Hydrolyse von THON. Das Intermediat 2-(Enamin)-glutarat ist fest am positiv geladenen Boden der aktiven Tasche fixiert, so dass die finale Hydrolyse nur von der *Re*-Seite des Substrats erfolgen kann. Zum Vergleich ist Carbamoylaspartat, das Produkt der Ec.DHO-Reaktion, gezeigt (gelb).

Die Protonierung des C2-Kohlenstoffatoms des 2-(Enamin)-glutarats ist entscheidend für die Enantioselektivität der Reaktion. Unserem Bindungsmodell zufolge ist die Addition eines Protons von der *Si*-Seite unwahrscheinlich, da es dort weder geeignete Protonendonoren noch Wassermoleküle gibt. Es scheint daher, als würde die Enantioselektivität der Enamidase dadurch erreicht werden, dass die Bildung von (*R*)-2-(Formimino)-glutarat verhindert wird.

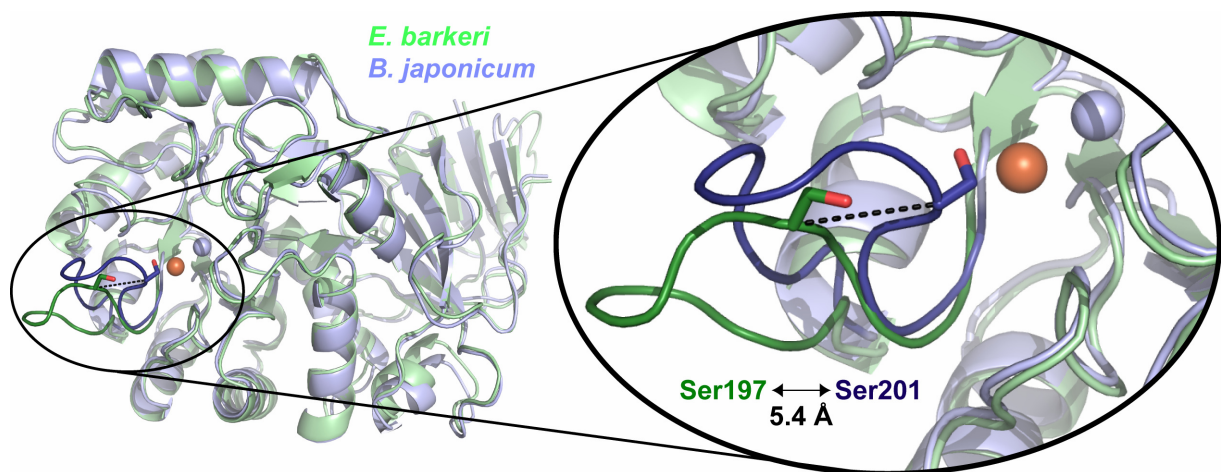
### 5.1.4 Ausblick

Einen wertvollen Beitrag zur Überprüfung des Substratbindungsmodus, der aus der Kristallstruktur der *E. barkeri*-Enamidase abgeleitet wurde, und des sich daraus ergebenden enantioselektiven Reaktionsmechanismus könnte die ortspezifische Mutagenese der entsprechenden Aminosäurereste und die anschließende funktionelle und strukturelle Charakterisierung der Enamidase-Mutanten leisten. Ferner wurden zwischenzeitlich weitere Enamidasen aus an-



deren Organismen charakterisiert, die sich auf Grund ihrer enzymkinetischen Eigenschaften besser für Kokristallisations- und soaking-Experimente eignen könnten (Pierik, persönliche Mitteilung). Ein Beispiel dafür ist die Enamidase aus dem Stickstoff-fixierenden Bakterium *Bradyrhizobium japonicum*, das eine deutlich höhere spezifische Aktivität für die THON-Hydrolyse aufweist als das *E. barkeri*-Enzym (vgl. 4.2.1, [Bozarth 2005]). Die *B. japonicum*-Enamidase konnte mittlerweile erfolgreich kristallisiert und ihre Kristallstruktur durch Molekularen Ersatz aufgeklärt werden (vgl. 4.2.2). Die endgültige Verfeinerung und Analyse des Strukturmodells steht zum jetzigen Zeitpunkt noch aus [Kress *et al.* Manuskript in Vorbereitung].

Bei einem ersten Vergleich der Strukturmodelle der Enzyme aus *B. japonicum* und *E. barkeri* konnten die Bewegung eines Schleifenbereichs in der unmittelbaren Nähe des aktiven Zentrums beobachtet und ein konservierter Serin-Rest identifiziert werden, der sich im Vergleich zur *E. barkeri*-Enamidase um 5.4 Å in Richtung des Metallzentrums „bewegt“ (Abbildung 5-5). Ob dieser oder weitere Aminosäurereste direkt an der Katalyse teilnehmen oder lediglich dazu beitragen, das Substrat sterisch zu fixieren, wird die weitere Analyse der Struktur zeigen müssen.



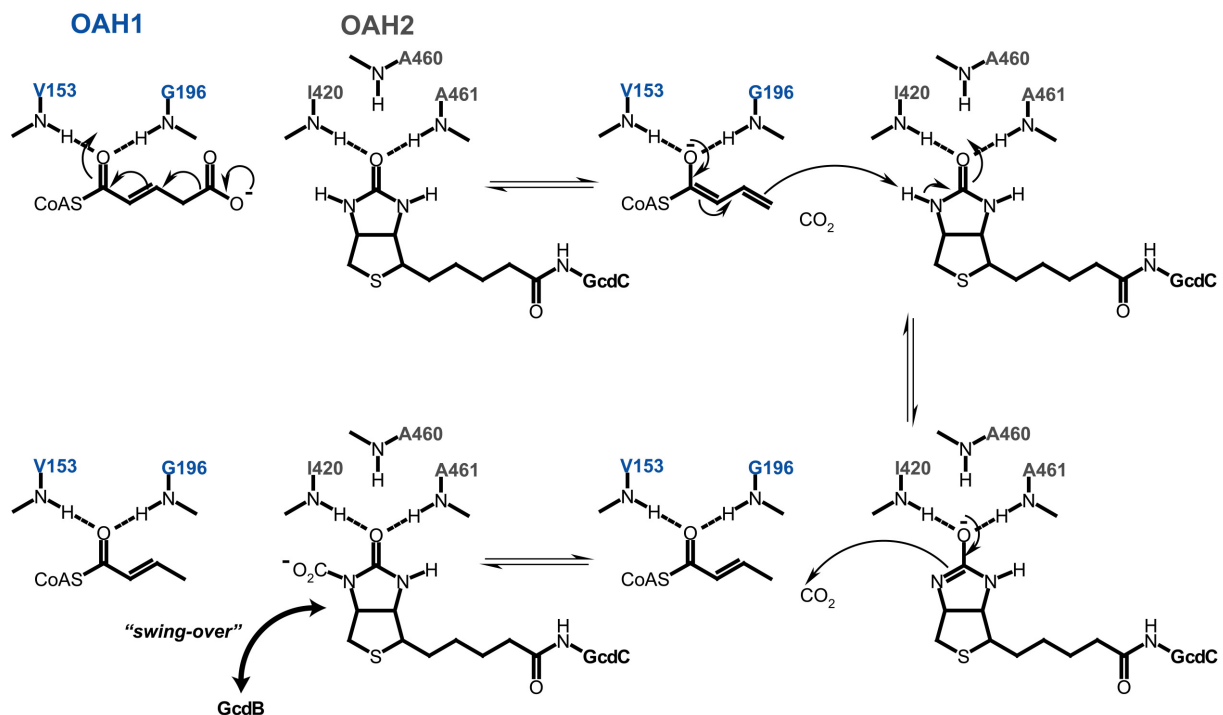
**Abbildung 5-5** Vorläufige Ergebnisse der Kristallstrukturanalyse der *Bradyrhizobium japonicum*-Enamidase Überlagerung der Strukturmodelle der Enzyme aus *B. japonicum* (hellblau) und *E. barkeri* (hellgrün). Die Eisen- (rot) und Zinkionen (hellblau) der *E. barkeri*-Enamidase sind als Kugeln dargestellt. Der offenbar bewegliche Schleifenbereich ist zusätzlich vergrößert gezeigt. Die Längenangabe bezieht sich auf die relative Position eines konservierten Serin-Restes, der sich in Richtung des Metallzentrums „bewegt“.

## 5.2 Glutaconyl-CoA-Decarboxylase

### 5.2.1 Mechanismus der Decarboxylase-Untereinheit

Abbildung 5-6 zeigt den hypothetischen Reaktionsverlauf des von GcdA katalysierten Carboxyltransfers von Glutaconyl-CoA auf Biotin. Die Grundvoraussetzung für diesen Mechanismus ist die Bindung der Carbonylgruppen-Sauerstoffatome der beiden Substrate an die Oxyanionenlöcher OAH1 und OAH2 und damit die Stabilisierung der Enolat-Zwischenstufen der Reaktion.

Mit Hilfe von NMR-spektroskopischen Methoden konnte gezeigt werden, dass der Carboxyltransfer direkt auf das N<sup>1</sup>-Stickstoffatom des Biotins erfolgt [Berger *et al.* 1996] und Letzteres schon zu Beginn der Reaktion zumindest teilweise in seiner Enolatform vorliegt [Bendrat *et al.* 1990]. Daher war bereits vor der Verfügbarkeit bestätigender Kristallstrukturdaten vermutet worden, dass die Katalyse ähnlich wie bei der Enoyl-CoA-Dehydratase über Enolat-Zwischenstufen verläuft, die durch Oxyanionenlöcher stabilisiert werden. Diese Annahme wird durch die in der vorliegenden Arbeit präsentierten Kokristallstrukturen und das Strukturmodell für *A. fermentans*-GcdA bestätigt [Kress *et al.* 2009; Wendt *et al.* 2003].

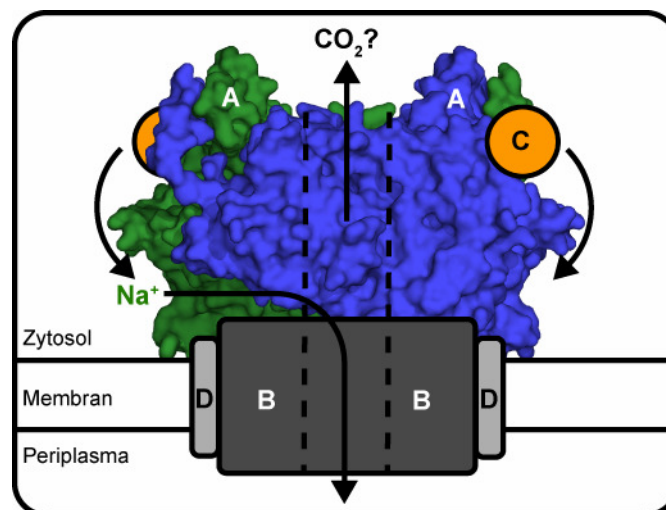


**Abbildung 5-6 Hypothetischer Mechanismus der Carboxyltransferase-Reaktion** Die Aminosäurereste der Oxyanionenlöcher sind entsprechend ihrer Zugehörigkeit zu den N- und C-terminalen Domänen der GcdA-Monomere blau bzw. grau gefärbt.

Da sich im Bereich des aktiven Zentrums keine sauren oder basischen Aminosäurereste befinden, die an der Katalyse beteiligt sein könnten, wird davon ausgegangen, dass die Reaktion mit der spontanen Decarboxylierung von Glutaconyl-CoA beginnt. Dass die Polarisation der Carbonyl-Gruppe des CoA-Esters dafür ausreicht, wird durch die in dieser Arbeit beschriebene Beobachtung gestützt, dass auch bei offensichtlich erfolgreicher Kokristallisation des Enzyms mit seinem Substrat Glutaconyl-CoA stets nur das Produkt Crotonyl-CoA in der Kristallstruktur nachweisbar ist.

### 5.2.2 Ein asymmetrisches Modell für Na<sup>+</sup>-pumpende Glutaconyl-CoA-Decarboxylasen

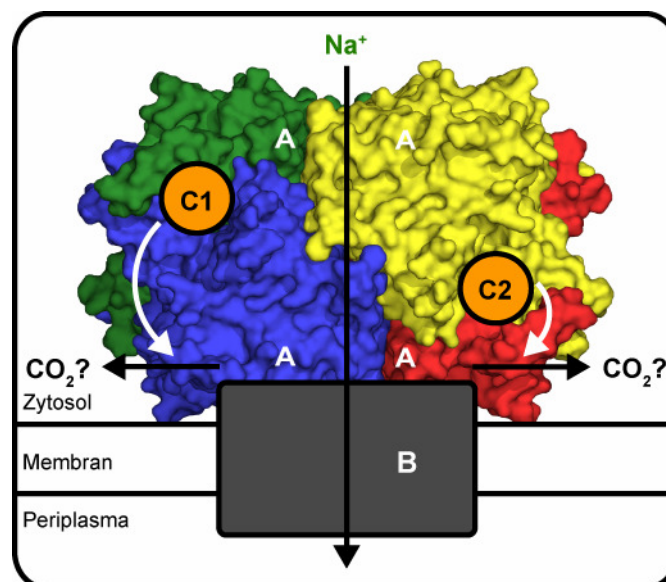
Wendt et al. postulierten 2003 auf der Grundlage von Gelfiltrationsexperimenten und der Analyse der Kristallstruktur der GcdA aus *A. fermentans* eine dimere Zusammensetzung für den intakten Glutaconyl-CoA-Decarboxylase-Komplex [Wendt *et al.* 2003]. Dabei wurde das die aktiven Zentren enthaltende GcdA-Dimer aus den Proteinketten A und B (vgl. 4.3.7) so mit der Na<sup>+</sup>-Pumpen-Untereinheit GcdB verbunden, wie in Abbildung 5-7 gezeigt.



**Abbildung 5-7 Aus der Dimerstruktur abgeleitetes symmetrisches Modell der Gcd** Die GcdA-Untereinheiten sind in der Oberflächendarstellung gezeigt. Das GcdA-Dimer hat die gleiche Orientierung und Farbe wie in Abbildung 4-17. Die Untereinheiten B, C und D sind jeweils dunkelgrau, orange und hellgrau eingefärbt. Die hypothetische Bewegung der GcdC-Untereinheiten ist durch gebogene Pfeile angedeutet (Abbildung nach [Wendt *et al.* 2003]).



Wahrscheinlicher ist jedoch, dass es sich bei GcdA um ein tetrameres Protein handelt, wie in 4.3.9 ausführlich dargelegt wurde: Sowohl die Volumenausschlusschromatographie als auch die Ergebnisse einer eingehenden Kristallstrukturanalyse deuten auf die Bildung stabiler GcdA-Tetramere hin. Diese Annahme wird außerdem durch die Untersuchung von GcdA-Lösungen mit Hilfe von Röntgenkleinwinkelstreuung (engl.: *small angle X-ray scattering*, SAXS, [Lipfert & Doniach 2007; Putnam *et al.* 2007]) bestätigt: Die Auswertung eines aus diesen Messungen erstellten Guinier-Plots ergab eine molekulare Masse von 245 kDa für die GcdA-Partikel in der Probenlösung [Pieles 2009].

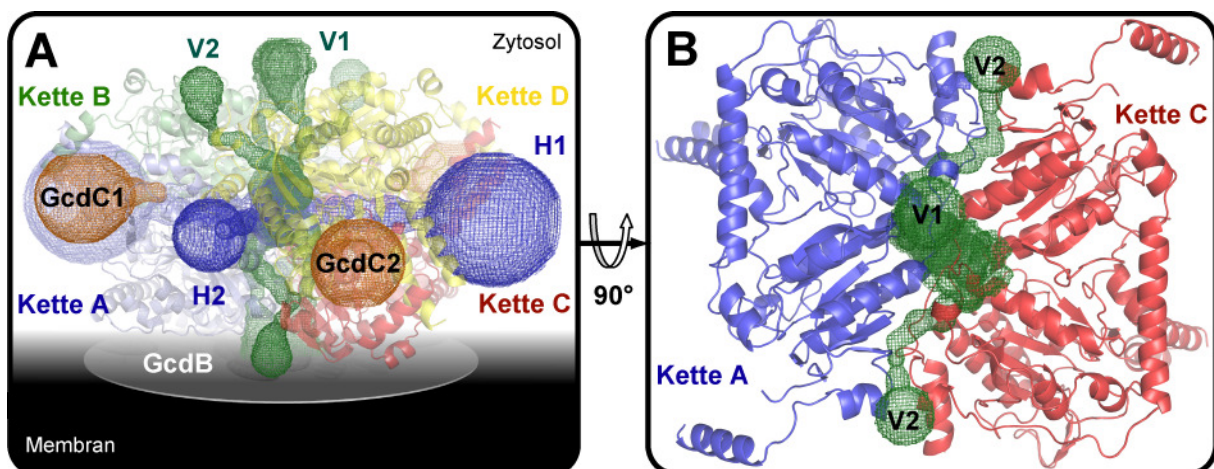


**Abbildung 5-8 Asymmetrisches Modell der Gcd auf Grundlage der Tetramerstruktur** Die GcdA-Untereinheiten sind in der Oberflächendarstellung gezeigt. Das GcdA-Tetramer hat die gleiche Orientierung und Farbe wie in Abbildung 4-20 (rechte Seite). Die Untereinheiten B und C1/C2 sind dunkelgrau und orange eingefärbt. Die hypothetische Bewegung der GcdC-Untereinheiten ist durch gebogene Pfeile angedeutet.

Auf Grund der intrinsischen Stabilität der GcdA-Tetramere und der Existenz von zwei verschiedenen Biotin-Carrier-Untereinheiten, GcdC<sub>1</sub> und GcdC<sub>2</sub>, wird daher ein alternatives, asymmetrisches Modell vorgeschlagen (Abbildung 5-8). Geht man davon aus, dass der Abstand der GcdC-Untereinheiten von der Membranuntereinheit GcdB ähnlich sein muss, gibt es nur eine mögliche Orientierung für das GcdA-Tetramer. Dabei kommt das AC-Dimer mit

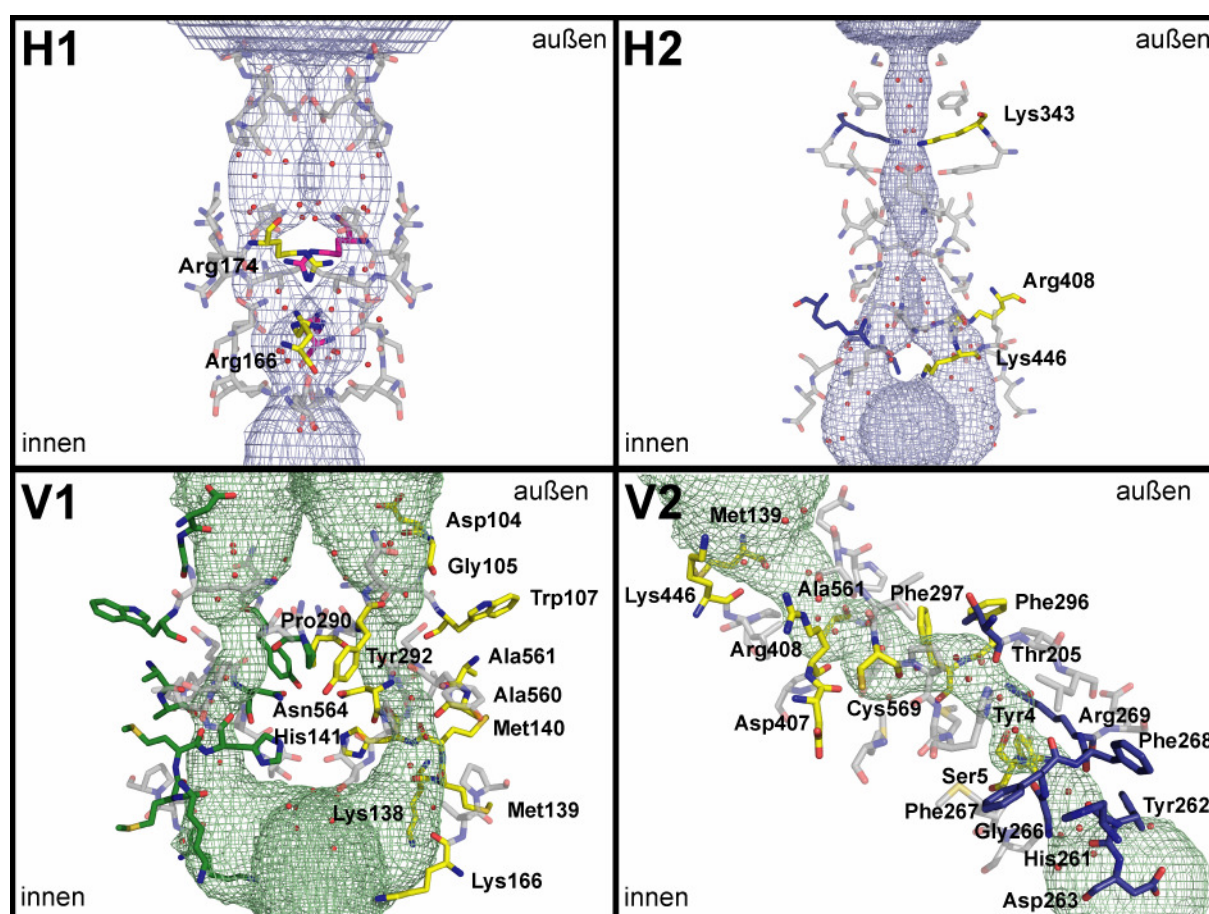
seiner flachen, annähernd rechteckigen und ca.  $50 \times 70 \text{ \AA}^2$  großen Oberfläche auf GcdB und der Membranoberfläche zum Liegen. Der Fluss von Natriumionen durch das GcdA-Tetramer wäre dann immer noch möglich – wie die folgenden Erläuterungen deutlich machen werden – oder könnte sogar von allosterisch induzierten Veränderungen der Quartärstruktur abhängig sein.

Eine Analyse des GcdA-Tetramers mit Hilfe des MolAxis-Servers [Yaffe *et al.* 2008] führte zur Entdeckung eines ausgeprägten sphärischen Hohlraums in seinem Zentrum, der mit der Oberfläche des Proteins durch ein Netzwerk von Kanälen verbunden ist. Durch die anschließende Untersuchung mit den Programmen MOLE [Petrek *et al.* 2007] und CAVER [Petrek *et al.* 2006] konnten vier verschiedene Arten von Lösungsmittel-gefüllten Tunneln identifiziert werden, die von dieser inneren Hohlkugel ausgehen (Abbildung 5-9A). Entsprechend ihrer relativen Orientierung zur Membranebene wurden die Kanäle als horizontale Tunnel H1 und H2 bzw. transversale Tunnel V1 und V2 benannt. Eine vergrößerte Darstellung der einzelnen Kanäle ist in Abbildung 5-10 gezeigt.



**Abbildung 5-9 Das Netzwerk Lösungsmittel-gefüllter Kanäle im Inneren des GcdA-Tetramers**  
 (A) Die Kanäle wurden mit dem Programm CAVER berechnet und werden entsprechend ihrer relativen Orientierung zur Ebene der hypothetischen Membran als horizontale Tunnel H1 und H2 (blau) bzw. transversale Tunnel V1 und V2 (grün) bezeichnet. Die sphärischen Enden der Tunnel sind Artefakte der visuellen Darstellung der CAVER-Berechnungen. Die in der Membran lokalisierte Untereinheit GcdB ist als graue Scheibe angedeutet. Das orange Geflecht zeigt die wahrscheinliche Bindestelle der Biotin-Carrier-Untereinheiten GcdC<sub>1</sub> und GcdC<sub>2</sub> sowie die zum OAH2 führenden Kanäle an. Orientierung und Farbgebung des GcdA-Tetramers entsprechen Abbildung 5-8. (B) Dargestellt ist die mit GcdB in Kontakt stehende Oberfläche, die durch die Ketten A und C gebildet wird (von GcdB aus betrachtet). Der Übersicht halber sind nur die vertikalen Tunnel V1 und V2 gezeigt. Von den  $\alpha$ -Helices, die die gezeigten Kanäle umgeben, sind nur diejenigen beschriftet, die sich in unmittelbarer Nähe der hypothetischen Membranoberfläche befinden.

Der ausgeprägteste Tunnel, H1, wurde bereits von Wendt et al. bei der Besprechung der Struktur der *A. fermentans*-GcdA erwähnt [Wendt et al. 2003]. Er wird jeweils von Aminosäureresten der Proteinketten A(B) und C(D) aufgebaut und verläuft entlang der C<sub>2</sub>-Symmetrieachse der AB- bzw. CD-Dimere. H1 enthält einen positiv geladenen „Flaschenhals“, der aus den Resten Arg166A(C), Arg166B(D), Arg174A(C) und Arg174B(D) besteht und daher ungeeignet für den Durchtritt von Na<sup>+</sup>-Ionen ist. In ähnlicher Weise ist durch die positiv geladenen Seitenketten von Arg408A(B), Arg408D(C), Lys343A(B), Lys343D(C), Lys446A(B) und Lys446D(C) der Zugang zu H2 blockiert.



**Abbildung 5-10** Vergrößerte Darstellung der vier verschiedenen Arten von horizontalen (H) und vertikalen (V) Kanälen im Inneren des GcdA-Tetramers. Aminosäurereste in einem Abstand von 5 Å vom Zentrum des jeweiligen Tunnels sind in grau dargestellt. Die „Flaschenhals-Reste“ in den horizontalen Tunneln und die Reste, welche die zwischenzeitliche Koordination monovalenter Kationen in den vertikalen Tunneln ermöglichen sollen, sind in den Farben der jeweiligen Proteinkette gezeigt (vgl. Abbildung 5-9). Im Falle von H1, V1 und V2 ist der Einfachheit halber nur die eine Hälfte der Aminosäurereste beschriftet. Wassermoleküle sind als rote Kugeln dargestellt. Die Abbildung wurde mit Hilfe des CAVER-Plugins für das Programm Pymol erstellt.

Im Vergleich dazu scheinen die vertikalen Kanäle trotz ihrer kleineren relativen Durchmesser eher für die Translokation von Kationen geeignet zu sein: Die inneren Oberflächen von V1 und V2 sind ausgekleidet mit den Carbonyl-Sauerstoffatomen der entsprechenden Aminosäurereste, die die vorübergehende Koordination monovalenter Kationen erlauben könnten. Darüber hinaus deuten die C-Termini der Helices, die V1 und V2 umgeben ( $\alpha 4$ ,  $\alpha 5$ ,  $\alpha 6$ ,  $\alpha 7$ ,  $\alpha 9$ ,  $\alpha 10$ ,  $\alpha 12$ ,  $\alpha 14$  und  $\alpha 19$  der jeweiligen Ketten), in Richtung der Kanäle und könnten daher die Translokation von Kationen durch Dipol-Ionen-Interaktion stabilisieren, wie es für den kaliumselektiven Ionenkanal KcsA und ähnliche Kanäle bereits beobachtet wurde (Abbildung 5-9B) [Doyle *et al.* 1998; Gouaux 1998].

Im hier vorgeschlagenen asymmetrischen Modell unterscheiden sich die Positionen der potentiellen Bindestellen für die Biotin-Carrier-Untereinheiten um ca. 12 Å relativ zur GcdA-GcdB-Kontaktfläche, an der die Decarboxylierung des Kosubstrats Biotin stattfindet (Abbildung 5-9A). Die Asymmetrie des Komplexes könnte somit erklären, warum zwei GcdC-Untereinheiten mit unterschiedlich langen Linkerbereichen zwischen der N-terminalen Ankerregion und der C-terminalen Biotin-Carrier-Domäne Bestandteil des Komplexes sind.

### 5.2.3 Ausblick

Die Glutaconyl-CoA-Decarboxylase wäre nicht der erste asymmetrische Enzymkomplex, der die Katalyse einer chemischen Reaktion mit der Translokation von Ionen koppelt: Die Interaktion des extramembranären  $\alpha_3\beta_3$ -Teilkomplexes in F-Typ-ATPasen mit der  $\gamma$ -Untereinheit und den Untereinheiten (b, b') des zweiten Stators weist ebenfalls eine gebrochene Symmetrie auf [Pedersen 2007]. Darüber hinaus konnte nachgewiesen werden, dass das trimere AcrB, ein Mitglied der RND-Familie, asymmetrische Konformationen annimmt, die abhängig sind von der Transporttätigkeit des Enzyms [Pos 2009]. Letztlich wird natürlich erst die Aufklärung und Analyse der Struktur des intakten Gcd-Komplexes zweifelsfrei bestätigen können, dass diese Natriumpumpe tatsächlich asymmetrische Komplexe ähnlich denen der F-Typ-ATPase bildet.

## 6 Literaturverzeichnis

Abendroth, J., Niefind, K. & Schomburg, D. (2002) X-ray structure of a dihydropyrimidinase from *Thermus* sp. at 1.3 Å resolution. *J Mol Biol* **320**, 143-56.

Abrahams, J. P. (1997) Bias reduction in phase refinement by modified interference functions: introducing the gamma correction. *Acta Crystallogr D Biol Crystallogr* **53**, 371-6.

Aimin, L., Tingfeng, L. & Rong, F. (2007) Amidohydrolase Superfamily. *ENCYCLOPEDIA OF LIFE SCIENCES John Wiley & Sons Ltd, Chichester*.

Alhapel, A. (2005) Nicotinatfermentation in *Eubacterium barkeri*: Klonierung des kompletten Nicotinat-Locus und Charakterisierung der 2-(Hydroxymethyl)glutarat-Dehydrogenase und Enamidase. *Dissertation, Fachbereich Biologie, Universität Marburg*.

Alhapel, A., Darley, D. J., Wagener, N., Eckel, E., Elsner, N. & Pierik, A. J. (2006) Molecular and functional analysis of nicotinate catabolism in *Eubacterium barkeri*. *Proc Natl Acad Sci USA* **103**, 12341-6.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **98**, 10037-41.

Banner, D. W., Bloomer, A., Petsko, G. A., Phillips, D. C. & Wilson, I. A. (1976) Atomic coordinates for triose phosphate isomerase from chicken muscle. *Biochem Biophys Res Commun* **72**, 146-55.

Barker, H. A. (1961) Fermentations of nitrogenous organic compounds. In *The bacteria* (Gun-salus G, ed<sup>eds</sup>), pp. 151–207. Academic Press, New York.

Barker, H. A. (1981) Amino acid degradation by anaerobic bacteria. *Annu Rev Biochem* **50**, 23-40.

Beatrix, B., Bendrat, K., Rospert, S. & Buckel, W. (1990) The biotin-dependent sodium ion pump glutaconyl-CoA decarboxylase from *Fusobacterium nucleatum* (subsp. *nucleatum*). Comparison with the glutaconyl-CoA decarboxylases from gram-positive bacteria. *Arch Microbiol* **154**, 362-9.

Becher, B., Muller, V. & Gottschalk, G. (1992) N5-methyl-tetrahydromethanopterin:coenzyme M methyltransferase of *Methanosarcina* strain Go1 is an Na(+)-translocating membrane protein. *J Bacteriol* **174**, 7656-60.

Behrman, E. J. & Stanier, R. Y. (1957) The bacterial oxidation of nicotinic acid. *J Biol Chem* **228**, 923-45.

- Bendrat, K., Berger, S., Buckel, W., Etzel, W. A. & Rohm, K. H. (1990) Carbon-13 labelled biotin--a new probe for the study of enzyme catalyzed carboxylation and decarboxylation reactions. *FEBS Lett* **277**, 156-8.
- Bennett, J. & Scott, K. J. (1971) Quantitative staining of fraction I protein in polyacrylamide gels using Coomassie brilliant blue. *Anal Biochem* **43**, 173-82.
- Benning, M. M., Haller, T., Gerlt, J. A. & Holden, H. M. (2000) New reactions in the crotonase superfamily: structure of methylmalonyl CoA decarboxylase from *Escherichia coli*. *Biochemistry* **39**, 4630-9.
- Benning, M. M., Kuo, J. M., Raushel, F. M. & Holden, H. M. (1995) Three-dimensional structure of the binuclear metal center of phosphotriesterase. *Biochemistry* **34**, 7973-8.
- Benning, M. M., Taylor, K. L., Liu, R. Q., Yang, G., Xiang, H., Wesenberg, G., Dunaway-Mariano, D. & Holden, H. M. (1996) Structure of 4-chlorobenzoyl coenzyme A dehalogenase determined to 1.8 Å resolution: an enzyme catalyst generated via adaptive mutation. *Biochemistry* **35**, 8103-9.
- Berger, S., Braune, A., Buckel, W., Härtel, U. & Lee, M. L. (1996) Enzyme catalysed formation of carboxybiotin as proved by the measurement of  $^{15}\text{N}$ ,  $^{13}\text{C}$  and  $^{13}\text{C}$ ,  $^{13}\text{C}$  spin-spin coupling. *Angew Chem Int Ed Engl* **35**, 2132-3.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res* **28**, 235-42.
- Bilder, P., Lightle, S., Bainbridge, G., Ohren, J., Finzel, B., Sun, F., Holley, S., Al-Kassim, L., Spessard, C., Melnick, M., Newcomer, M. & Waldrop, G. L. (2006) The structure of the carboxyltransferase component of acetyl-coA carboxylase reveals a zinc-binding motif unique to the bacterial enzyme. *Biochemistry* **45**, 1712-22.
- Boiangiu, C. D., Jayamani, E., Brugel, D., Herrmann, G., Kim, J., Forzi, L., Hedderich, R., Vgenopoulou, I., Pierik, A. J., Steuber, J. & Buckel, W. (2005) Sodium ion pumps and hydrogen production in glutamate fermenting anaerobic bacteria. *J Mol Microbiol Biotechnol* **10**, 105-19.
- Bourenkov, G. P. & Popov, A. N. (2006) A quantitative approach to data-collection strategies. *Acta Crystallogr D Biol Crystallogr* **62**, 58-64.



- Bozarth, A. (2005) Examination of a Common Nicotinate Metabolism in Proteobacteria. *Diplomarbeit, Fachbereich Biologie, Philipps-Universität Marburg*.
- Braune, A., Bendrat, K., Rospert, S. & Buckel, W. (1999) The sodium ion translocating glutacetyl-CoA decarboxylase from *Acidaminococcus fermentans*: cloning and function of the genes forming a second operon. *Mol Microbiol* **31**, 473-87.
- Bremer, E. & Krämer, R. (2000) Coping with osmotic challenges: Osmoregulation through accumulation and release of compatible solutes in bacteria. In *Bacterial stress responses* (Storz, G. & Hengge-Aronis, R.), ASM Press, Washington, D.C, 79-97.
- Brenner, S. E., Chothia, C. & Hubbard, T. J. P. (1997) Population statistics of protein structures: lessons from structural classifications. *Current Opinion in Structural Biology* **7**, 369-76.
- Bricogne, G. (1974) Geometric sources of redundancy in intensity data and their use for phase determination. *Acta Crystallographica Section A* **30**, 395-405.
- Brünger, A. (1993) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallographica Section D* **49**, 24-36.
- Brünger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-5.
- Buchbinder, J. L., Stephenson, R. C., Dresser, M. J., Pitera, J. W., Scanlan, T. S. & Fletterick, R. J. (1998) Biochemical characterization and crystallographic structure of an *Escherichia coli* protein from the phosphotriesterase gene family. *Biochemistry* **37**, 5096-106.
- Buckel, W. (1980) Analysis of the fermentation pathways of clostridia using double labelled glutamate. *Arch Microbiol* **127**, 167-9.
- Buckel, W. (1980) The reversible dehydration of (R)-2-hydroxyglutarate to (E)-glutaconate. *Eur J Biochem* **106**, 439-47.
- Buckel, W. (2001a) Sodium ion-translocating decarboxylases. *Biochim Biophys Acta* **1505**, 15-27.
- Buckel, W. (2001b) Unusual enzymes involved in five pathways of glutamate fermentation. *Appl Microbiol Biotechnol* **57**, 263-73.
- Buckel, W. & Barker, H. A. (1974) Two pathways of glutamate fermentation by anaerobic bacteria. *J Bacteriol* **117**, 1248-60.
- Buckel, W. & Golding, B. T. (1998) Radical species in the catalytic pathways of enzymes from anaerobes. *FEMS Microbiology Reviews* **22**, 523-41.

- Buckel, W., Janssen, P., Schuhmann, A., Eikmanns, U., Messner, P., Sleytr, U. & Liesack, W. (1994) *Clostridium viride* sp. nov., a strictly anaerobic bacterium using 5-aminovalerate as growth substrate, previously assigned to *Clostridium aminovalericum*. *Archives of Microbiology* **162**, 387-94.
- Buckel, W. & Semmler, R. (1982) A biotin-dependent sodium pump: glutaconyl-CoA decarboxylase from *Acidaminococcus fermentans*. *FEBS Lett* **148**, 35-8.
- Buckel, W. & Semmler, R. (1983) Purification, characterisation and reconstitution of glutaconyl-CoA decarboxylase, a biotin-dependent sodium pump from anaerobic bacteria. *Eur J Biochem* **136**, 427-34.
- Castanie-Cornet, M. P. & Foster, J. W. (2001) *Escherichia coli* acid resistance: cAMP receptor protein and a 20 bp cis-acting sequence control pH and stationary phase expression of the *gadA* and *gadBC* glutamate decarboxylase genes. *Microbiology* **147**, 709-15.
- Castanie-Cornet, M. P., Penfound, T. A., Smith, D., Elliott, J. F. & Foster, J. W. (1999) Control of acid resistance in *Escherichia coli*. *J Bacteriol* **181**, 3525-35.
- CCP4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Cryst D* **50**, 760-3.
- Collins, M. D., Lawson, P. A., Willems, A., Cordoba, J. J., Fernandez-Garayzabal, J., Garcia, P., Cai, J., Hippe, H. & Farrow, J. A. (1994) The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol* **44**, 812-26.
- Cramer, W. A. & Knaff, D. B. (1990) *Energy Transduction in Biological Membranes: A Textbook of Bioenergetics* Springer-Verlag, Berlin.
- Cuff, A. L., Sillitoe, I., Lewis, T., Redfern, O. C., Garratt, R., Thornton, J. & Orengo, C. A. (2009) The CATH classification revisited--architectures reviewed and new ways to characterize structural divergence in superfamilies. *Nucleic Acids Res* **37**, D310-4.
- Damborsky, J., Petrek, M., Banas, P. & Otyepka, M. (2007) Identification of tunnels in proteins, nucleic acids, inorganic materials and molecular ensembles. *Biotechnol J* **2**, 62-7.
- Davies, D. R. & Segal, D. M. (1971) Protein crystallization: Micro techniques involving vapor diffusion *Methods in Enzymology* **22**, 266-9
- Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J. & Swanson, R. V. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353-8.
- DeLano, W. L. (2002) *The PyMOL molecular graphics system*. In ed^eds). DeLano Scientific LLC, <http://www.pymol.org>, San Carlos, CA, USA.



- Deo, R. C., Schmidt, E. F., Elhabazi, A., Togashi, H., Burley, S. K. & Strittmatter, S. M. (2004) Structural bases for CRMP function in plexin-dependent semaphorin3A signaling. *Embo J* **23**, 9-22.
- Di Berardino, M. & Dimroth, P. (1996) Aspartate 203 of the oxaloacetate decarboxylase beta-subunit catalyses both the chemical and vectorial reaction of the Na<sup>+</sup> pump. *Embo J* **15**, 1842-9.
- Diacovich, L., Mitchell, D. L., Pham, H., Gago, G., Melgar, M. M., Khosla, C., Gramajo, H. & Tsai, S. C. (2004) Crystal structure of the beta-subunit of acyl-CoA carboxylase: structure-based engineering of substrate specificity. *Biochemistry* **43**, 14027-36.
- Diederichs, K. & Karplus, P. A. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* **4**, 269-75.
- Dimroth, P. (1982) The generation of an electrochemical gradient of sodium ions upon decarboxylation of oxaloacetate by the membrane-bound and Na<sup>+</sup>-activated oxaloacetate decarboxylase from *Klebsiella aerogenes*. *Eur J Biochem* **121**, 443-9.
- Dimroth, P. (1990) Bacterial energy transductions coupled to sodium ions. *Res Microbiol* **141**, 332-6.
- Dimroth, P. (1997) Primary sodium ion translocating enzymes. *Biochim Biophys Acta* **1318**, 11-51.
- Dimroth, P. & Hilbi, H. (1997) Enzymic and genetic basis for bacterial growth on malonate. *Mol Microbiol* **25**, 3-10.
- Dimroth, P., Jockel, P. & Schmid, M. (2001) Coupling mechanism of the oxaloacetate decarboxylase Na<sup>+</sup> pump. *Biochim Biophys Acta* **1505**, 1-14.
- Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. & MacKinnon, R. (1998) The structure of the potassium channel: molecular basis of K<sup>+</sup> conduction and selectivity. *Science* **280**, 69-77.
- Ducruix, A. & Giegé, R. (1999) *Crystallization of nucleic acids and proteins: a practical approach (2nd edition)*. Oxford University Press, Oxford, New York.
- Eggerer, H. (1985) Completion of the degradation scheme for nicotinic acid by *Clostridium barkeri*. *Curr Top Cell Regul* **26**, 411-8.
- Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Cryst D* **60**, 2126-32.

- Engel, C. K., Mathieu, M., Zeelen, J. P., Hiltunen, J. K. & Wierenga, R. K. (1996) Crystal structure of enoyl-coenzyme A (CoA) hydratase at 2.5 angstroms resolution: a spiral fold defines the CoA-binding pocket. *EMBO J* **15**, 5135-45.
- Engh, R. A. & Huber, R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallographica Section A* **47**, 392-400.
- Ensign, J. C. & Rittenberg, S. C. (1964) The pathway of nicotinic acid oxidation by a *Bacillus* species. *J Biol Chem* **239**, 2285-91.
- Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72-82.
- Evans, P. & McCoy, A. (2008) An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr* **64**, 1-10.
- Falb, M., Pfeiffer, F., Palm, P., Rodewald, K., Hickmann, V., Tittor, J. & Oesterhelt, D. (2005) Living with two extremes: conclusions from the genome sequence of *Natronomonas pharaonis*. *Genome Res* **15**, 1336-43.
- Ferreira, F. M., Mendoza-Hernandez, G., Castaneda-Bueno, M., Aparicio, R., Fischer, H., Calcagno, M. L. & Oliva, G. (2006) Structural analysis of *N*-acetylglucosamine-6-phosphate deacetylase apoenzyme from *Escherichia coli*. *J Mol Biol* **359**, 308-21.
- Finn, R. D., Tate, J., Mistry, J., Coghill, P. C., Sammut, S. J., Hotz, H. R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L. & Bateman, A. (2008) The Pfam protein families database. *Nucleic Acids Res* **36**, D281-8.
- Frishman, D. & Argos, P. (1995) Knowledge-based protein secondary structure assignment. *Proteins* **23**, 566-79.
- Garcia-Ruiz, J. M., Gonzalez-Ramirez, L. A., Gavira, J. A. & Otalora, F. (2002) Granada Crystallisation Box: a new device for protein crystallisation by counter-diffusion techniques. *Acta Crystallogr D Biol Crystallogr* **58**, 1638-42.
- Gartner, P., Weiss, D. S., Harms, U. & Thauer, R. K. (1994) N5-methyltetrahydromethanopterin:coenzyme M methyltransferase from *Methanobacterium thermoautotrophicum*. Catalytic mechanism and sodium ion dependence. *Eur J Biochem* **226**, 465-72.
- Giorgetti, A., Raimondo, D., Miele, A. E. & Tramontano, A. (2005) Evaluating the usefulness of protein structure models for molecular replacement. *Bioinformatics* **21 Suppl 2**, ii72-6.
- Gouaux, E. (1998) Single potassium ion seeks open channel for transmembrane travels: tales from the KcsA structure. *Structure* **6**, 1221-6.

- Hall, P. R., Wang, Y. F., Rivera-Hainaj, R. E., Zheng, X., Pustai-Carey, M., Carey, P. R. & Yee, V. C. (2003) Transcarboxylase 12S crystal structure: hexamer assembly and substrate binding to a multienzyme core. *EMBO J* **22**, 2334-47.
- Hall, R. S., Brown, S., Fedorov, A. A., Fedorov, E. V., Xu, C., Babbitt, P. C., Almo, S. C. & Raushel, F. M. (2007) Structural diversity within the mononuclear and binuclear active sites of *N*-acetyl-D-glucosamine-6-phosphate deacetylase. *Biochemistry* **46**, 7953-62.
- Haller, T., Buckel, T., Retey, J. & Gerlt, J. A. (2000) Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by *Escherichia coli*. *Biochemistry* **39**, 4622-9.
- Hamed, R. B., Batchelar, E. T., Clifton, I. J. & Schofield, C. J. (2008) Mechanisms and structures of crotonase superfamily enzymes--how nature controls enolate and oxyanion reactivity. *Cell Mol Life Sci* **65**, 2507-27.
- Hans, M., Buckel, W. & Bill, E. (2000) The iron-sulfur clusters in 2-hydroxyglutaryl-CoA dehydratase from *Acidaminococcus fermentans*. Biochemical and spectroscopic investigations. *Eur J Biochem* **267**, 7082-93.
- Hans, M., Sievers, J., Muller, U., Bill, E., Vorholt, J. A., Linder, D. & Buckel, W. (1999) 2-Hydroxyglutaryl-CoA dehydratase from *Clostridium symbiosum*. *Eur J Biochem* **265**, 404-14.
- Harada, Y., Lifchitz, A., Berthou, J. & Jolles, P. (1981) A translation function combining packing and diffraction information: an application to lysozyme (high-temperature form). *Acta Crystallographica Section A* **37**, 398-406.
- Harary, I. (1957) Bacterial fermentation of nicotinic acid. I. End products. *J Biol Chem* **227**, 815-22.
- Hardman, J. K. & Stadtman, T. C. (1960) Metabolism of omega-amino acids. I. Fermentation of gamma-aminobutyric acid by *Clostridium aminobutyricum* n. sp. *J Bacteriol* **79**, 544-8.
- Harker, D. (1956) The determination of the phases of the structure factors of non-centrosymmetric crystals by the method of double isomorphous replacement. *Acta Crystallographica* **9**, 1-9.
- Harp, J. M., Timm, D. E. & Bunick, G. J. (1998) Macromolecular crystal annealing: overcoming increased mosaicity associated with cryocrystallography. *Acta Crystallogr D Biol Crystallogr* **54**, 622-8.
- Hase, C. C., Fedorova, N. D., Galperin, M. Y. & Dibrov, P. A. (2001) Sodium ion cycle in bacterial pathogens: evidence from cross-genome comparisons. *Microbiol Mol Biol Rev* **65**, 353-70.

- Hayashi, M., Nakayama, Y. & Unemoto, T. (2001) Recent progress in the Na(+)-translocating NADH-quinone reductase from the marine *Vibrio alginolyticus*. *Biochim Biophys Acta* **1505**, 37-44.
- Heinig, M. & Frishman, D. (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res* **32**, W500-2.
- Hendrickson, W. A. & Konnert, J. H. (1980) Diffraction Analysis of Motion in Proteins. *Biophys J* **32**, 645-7.
- Henrick, K. & Thornton, J. M. (1998) PQS: a protein quaternary structure file server. *Trends Biochem Sci* **23**, 358-61.
- Hermann, J. C., Marti-Arbona, R., Fedorov, A. A., Fedorov, E., Almo, S. C., Shoichet, B. K. & Raushel, F. M. (2007) Structure-based activity prediction for an enzyme of unknown function. *Nature* **448**, 775-9.
- Hilpert, W., Schink, B. & Dimroth, P. (1984) Life by a new decarboxylation-dependent energy conservation mechanism with Na as coupling ion. *EMBO J* **3**, 1665-70.
- Holm, L. & Sander, C. (1997) An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. *Proteins* **28**, 72-82.
- Imhoff-Stuckle, D. & Pfennig, N. (1983) Isolation and characterization of a nicotinic acid-degrading sulfate reducing bacterium, *Desulfococcus niacini* sp. nov. *Arch Microbiol* **136**, 194-8.
- Ireton, G. C., McDermott, G., Black, M. E. & Stoddard, B. L. (2002) The structure of *Escherichia coli* cytosine deaminase. *J Mol Biol* **315**, 687-97.
- Jabri, E., Carr, M. B., Hausinger, R. P. & Karplus, P. A. (1995) The crystal structure of urease from *Klebsiella aerogenes*. *Science* **268**, 998-1004.
- Jancarik, J. & Kim, S. H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of Applied Crystallography* **24**, 409-11.
- Jozic, D., Kaiser, J. T., Huber, R., Bode, W. & Maskos, K. (2003) X-ray structure of isoaspartyl dipeptidase from *E. coli*: a dinuclear zinc peptidase evolved from amidohydrolases. *J Mol Biol* **332**, 243-56.
- Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Otsuka, R., Nakazawa, H., Takamiya, M., Kato, Y., Yoshizawa, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Masuda, S., Yanagii, M., Nishimura, M., Yamagishi, A., Oshima, T. & Kikuchi, H. (2001) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7. *DNA Res* **8**, 123-40.

- Kitts, C. L., Lapointe, J. P., Lam, V. T. & Ludwig, R. A. (1992) Elucidation of the complete *Azorhizobium* nicotinate catabolism pathway. *J Bacteriol* **174**, 7791-7.
- Kleywegt, G. (1996) Use of Non-crystallographic Symmetry in Protein Structure Refinement. *Acta Crystallographica Section D* **52**, 842-57.
- Kleywegt, G. J. & Jones, T. A. (1996) xdlMAPMAN and xdlDATAMAN - Programs for Re-formatting, Analysis and Manipulation of Biomacromolecular Electron-Density Maps and Reflection Data Sets. *Acta Crystallographica Section D* **52**, 826-8.
- Kleywegt, G. J. & Jones, T. A. (1998) Databases in protein crystallography. *Acta Crystallogr D Biol Crystallogr* **54**, 1119-31.
- Knowles, J. R. (1980) Enzyme-catalyzed phosphoryl transfer reactions. *Annu Rev Biochem* **49**, 877-919.
- Kogure, K. (1998) Bioenergetics of marine bacteria. *Curr Opin Biotechnol* **9**, 278-82.
- Krebs, W., Steuber, J., Gemperli, A. C. & Dimroth, P. (1999) Na<sup>+</sup> translocation by the NADH:ubiquinone oxidoreductase (complex I) from *Klebsiella pneumoniae*. *Mol Microbiol* **33**, 590-8.
- Kress, D., Alhapel, A., Pierik, A. J. & Essen, L. O. (2008) The crystal structure of enamidase: a bifunctional enzyme of the nicotinate catabolism. *J Mol Biol* **384**, 837-47.
- Kress, D., Bozarth, A., Pierik, A. J. & Essen, L. O. (Manuskript in Vorbereitung).
- Kress, D., Bruegel, D., Schall, I., Linder, D., Buckel, W. & Essen, L. O. (2009) An asymmetric model for Na<sup>+</sup>-translocating glutaconyl-CoA decarboxylases. *J Biol Chem*.
- Kriminski, S., Caylor, C. L., Nonato, M. C., Finkelstein, K. D. & Thorne, R. E. (2002) Flash-cooling and annealing of protein crystals. *Acta Crystallogr D Biol Crystallogr* **58**, 459-71.
- Krissinel, E. & Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst D* **60**, 2256-68.
- Krissinel, E. & Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallog sect D* **60**, 2256-68.
- Krissinel, E. & Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97.
- Krulwich, T. A. (1995) Alkaliphiles: 'basic' molecular problems of pH tolerance and bioenergetics. *Mol Microbiol* **15**, 403-10.

- Kulla, H. G. (1991) Enzymatic hydroxylations in industrial application. *Chimia* **45**, 81-5.
- Kung, H., Tsai, L. & Stadtman, T. C. (1971) Nicotinic acid metabolism. VIII. Tracer studies on the intermediary roles of  $\alpha$ -methyleneglutarate, methylitaconate, dimethylmaleate, and pyruvate. *J Biol Chem* **246**, 6444-51.
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5.
- Lamzin, V. S. & Wilson, K. S. (1993) Automated refinement of protein models. *Acta Crystallographica Section D* **49**, 129-47.
- Lamzin, V. S. & Wilson, K. S. (1997) Automated refinement for protein crystallography. *Methods Enzymol* **277**, 269-305.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & M., T. J. (1993) **PROCHECK**: a program to check the stereochemical quality of protein structures. *J Appl Chrystallogr* **26**, 283-91.
- Leslie, A. (1987) A reciprocal-space method for calculating a molecular envelope using the algorithm of B.C. Wang. *Acta Crystallographica Section A* **43**, 134-6.
- Leslie, A. G. (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**, 48-57.
- Leslie, A. G., Powell, H. R., Winter, G., Svensson, O., Spruce, D., McSweeney, S., Love, D., Kinder, S., Duke, E. & Nave, C. (2002) Automation of the collection and processing of X-ray diffraction data -- a generic approach. *Acta Crystallogr D Biol Crystallogr* **58**, 1924-8.
- Liao, D. I., Qian, J., Chisholm, D. A., Jordan, D. B. & Diner, B. A. (2000) Crystal structures of the photosystem II D1 C-terminal processing protease. *Nat Struct Biol* **7**, 749-53.
- Liaw, S. H., Chen, S. J., Ko, T. P., Hsu, C. S., Chen, C. J., Wang, A. H. & Tsai, Y. C. (2003) Crystal structure of D-aminoacylase from *Alcaligenes faecalis* DA1. A novel subset of amidohydrolases and insights into the enzyme mechanism. *J Biol Chem* **278**, 4957-62.
- Lipfert, J. & Doniach, S. (2007) Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu Rev Biophys Biomol Struct* **36**, 307-27.
- Locher, K. P., Hans, M., Yeh, A. P., Schmid, B., Buckel, W. & Rees, D. C. (2001) Crystal structure of the Acidaminococcus fermentans 2-hydroxyglutaryl-CoA dehydratase component A. *J Mol Biol* **307**, 297-308.
- Lohkamp, B., Andersen, B., Piskur, J. & Dobritsch, D. (2006) The crystal structures of dihydropyrimidinases reaffirm the close relationship between cyclic amidohydrolases and explain their substrate specificity. *J Biol Chem* **281**, 13762-76.

- Lunin, V. (1988) Use of the information on electron density distribution in macromolecules. *Acta Crystallographica Section A* **44**, 144-50.
- Martin, P. D., Purcarea, C., Zhang, P., Vaishnav, A., Sadecki, S., Guy-Evans, H. I., Evans, D. R. & Edwards, B. F. (2005) The crystal structure of a novel, latent dihydroorotase from *Aquifex aeolicus* at 1.7 Å resolution. *J Mol Biol* **348**, 535-47.
- Massa, W. (2007) *Kristallstrukturbestimmung* B.G. Teubner, Wiesbaden.
- Matthews, B. W. (1968) Solvent content of protein crystals. *J Mol Biol* **33**, 491-7.
- McPherson, A. (1999) *Crystallisation of Biological Macromolecules*. Cold Spring Harbor Laboratory Press.
- McPherson, A. (2003) *Introduction to macromolecular crystallography* John Wiley & Sons, Hoboken, New Jersey, USA.
- Michael, B. & Rainer, S. (2006) Chemie und Biologie Hand in Hand. Enzyme als Katalysatoren. *Chemie in unserer Zeit* **40**, 104-11.
- Mulkidjanian, A., Galperin, M., Makarova, K., Wolf, Y. & Koonin, E. (2008a) Evolutionary primacy of sodium bioenergetics. *Biology Direct* **3**, 13.
- Mulkidjanian, A. Y., Dibrov, P. & Galperin, M. Y. (2008b) The past and present of sodium energetics: May the sodium-motive force be with you. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1777**, 985-92.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst D* **53**, 240-55.
- Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**, 536-40.
- Naday, I., Westbrook, E. M., Westbrook, M. L., Travis, D. J., Stanton, M., Phillips, W. C., O'Mara, D. & Xie, J. (1994) Characterization and data collection on a direct-coupled CCD X-ray detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **348**, 635-40.
- Nanninga, H. J., Drent, W. J. & Gottschal, J. C. (1987) Fermentation of glutamate by *Selenomonas acidaminophila* sp. nov. *Archives of Microbiology* **147**, 152-7.
- Otwinowski, Z. & Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Volume 276: Macromolecular Crystallography, part A* (Carter CWJ & Sweet RM, eds), pp. 307-26. Academic Press, New York.

- Padan, E. & Schuldiner, S. (1993) Na<sup>+</sup>/H<sup>+</sup> antiporters, molecular devices that couple the Na<sup>+</sup> and H<sup>+</sup> circulation in cells. *J Bioenerg Biomembr* **25**, 647-69.
- Pedersen, P. L. (2007) Transport ATPases into the year 2008: a brief overview related to types, structures, functions and roles in health and disease. *J Bioenerg Biomembr* **39**, 349-55.
- Perrakis, A., Morris, R. & Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**, 458-63.
- Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. (1997) wARP: Improvement and Extension of Crystallographic Phases by Weighted Averaging of Multiple-Refined Dummy Atomic Models. *Acta Crystallographica Section D* **53**, 448-55.
- Petrek, M., Kosinova, P., Koca, J. & Otyepka, M. (2007) MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels. *Structure* **15**, 1357-63.
- Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J. & Damborsky, J. (2006) CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics* **7**, 316.
- Pfenninger-Li, X. D., Albracht, S. P., van Belzen, R. & Dimroth, P. (1996) NADH:ubiquinone oxidoreductase of *Vibrio alginolyticus*: purification, properties, and reconstitution of the Na<sup>+</sup> pump. *Biochemistry* **35**, 6233-42.
- Pieles, K. (2009) Interaktion zwischen dem Blaulichtrezeptor YtvA und dem Stressosom. *Diplomarbeit, Fachbereich Chemie, Philipps-Universität Marburg*.
- Pierik, A. J., Wolbert, R. B., Mutsaers, P. H., Hagen, W. R. & Veeger, C. (1992) Purification and biochemical characterization of a putative [6Fe-6S] prismane-cluster-containing protein from *Desulfovibrio vulgaris* (Hildenborough). *Eur J Biochem* **206**, 697-704.
- Plugge, C. M., van Leeuwen, J. M., Hummelen, T., Balk, M. & Stams, A. J. (2001) Elucidation of the pathways of catabolic glutamate conversion in three thermophilic anaerobic bacteria. *Arch Microbiol* **176**, 29-36.
- Ponstingl, H., Kabir, T. & M., T. J. (2003) Automatic inference of protein quaternary structure from crystals. *J Appl Chrystallogr* **36**, 1116-22.
- Pos, K. M. (2009) Drug transport mechanism of the AcrB efflux pump. *Biochim Biophys Acta* **1794**, 782-93.
- Pratt, J. W. (1976) F. Y. Edgeworth and R. A. Fisher on the Efficiency of Maximum Likelihood Estimation. *The Annals of Statistics* **4**, 501-14.



Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40**, 191-285.

Read, R. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallographica Section A* **42**, 140-9.

Rhodes, G. (2006) *Crystallography made crystal clear* Academic Press, Burlington, San Diego, London.

Rossmann, M. G. (1972) The Molecular Replacement Method. A collection of papers on the use of non-crystallographic symmetry. In ed^eds). Gordon & Breach, New York.

Rossmann, M. G. (1990) The molecular replacement method. *Acta Crystallogr A* **46 ( Pt 2)**, 73-82.

Rossmann, M. G. a. B., D.M. (1962) The Detection of Sub-Units Within the Crystallographic Asymmetric Unit. *Acta Crystallogr* **15**, 24-31.

Sander, C. & Schneider, R. (1991) Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9**, 56-68.

Schmidt, T. G. & Skerra, A. (2007) The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* **2**, 1528-35.

Schweiger, G., Dutscho, R. & Buckel, W. (1987) Purification of 2-hydroxyglutaryl-CoA dehydratase from *Acidaminococcus fermentans*. An iron-sulfur protein. *Eur J Biochem* **169**, 441-8.

Seibert, C. M. & Raushel, F. M. (2005) Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry* **44**, 6383-91.

She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P. M., Heikamp-de Jong, I., Jeffries, A. C., Kozera, C. J., Medina, N., Peng, X., Thi-Ngoc, H. P., Redder, P., Schenk, M. E., Theriault, C., Tolstrup, N., Charlebois, R. L., Doolittle, W. F., Duguet, M., Gaasterland, T., Garrett, R. A., Ragan, M. A., Sensen, C. W. & Van der Oost, J. (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc Natl Acad Sci U S A* **98**, 7835-40.

Skulachev, V. P. (1985) Membrane-linked energy transductions. Bioenergetic functions of sodium: H<sup>+</sup> is not unique as a coupling ion. *Eur J Biochem* **151**, 199-208.

Skulachev, V. P. (1988) *Membrane Bioenergetics* Springer-Verlag, Berlin.

- Skulachev, V. P. (1989) The sodium cycle: a novel type of bacterial energetics. *J Bioenerg Biomembr* **21**, 635-47.
- Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. & Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid. *Anal Biochem* **150**, 76-85.
- Speelmans, G., Poolman, B., Abee, T. & Konings, W. N. (1993) Energy transduction in the thermophilic anaerobic bacterium *Clostridium fervidus* is exclusively coupled to sodium ions. *Proc Natl Acad Sci U S A* **90**, 7975-9.
- Stadtman, E. R., Stadtman, T. C., Pastan, I. & Smith, L. D. (1972) *Clostridium barkeri* sp. n. *J Bacteriol* **110**, 758-60.
- Steuber, J., Schmid, C., Rufibach, M. & Dimroth, P. (2000) Na<sup>+</sup> translocation by complex I (NADH:quinone oxidoreductase) of *Escherichia coli*. *Mol Microbiol* **35**, 428-34.
- Stickland, L. H. (1935) Studies in the metabolism of the strict anaerobes (Genus *Clostridium*): The reduction of proline by *Cl. sporogenes*. *Biochem J* **29**, 288-90.
- Sussman, J. L., Harold W. Wyckoff, C. H. W. H. & Serge, N. T. (1985) Constrained-restrained least-squares (CORELS) refinement of proteins and nucleic acids. In *Methods in Enzymology* (ed^eds), pp. 271-303. Academic Press.
- Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F., Hiram, C., Nakamura, Y., Ogasawara, N., Kuhara, S. & Horikoshi, K. (2000) Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. *Nucleic Acids Res* **28**, 4317-31.
- Terwilliger, T. (1994) MAD phasing: Bayesian estimates of FA. *Acta Crystallographica Section D* **50**, 11-6.
- Terwilliger, T. (2004) SOLVE and RESOLVE: automated structure solution, density modification and model building. *J Synchrotron Radiat* **11**, 49-52.
- Terwilliger, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* **56**, 965-72.
- Terwilliger, T. C., Kim, S. H. & Eisenberg, D. (1987) Generalized method of determining heavy-atom positions using the difference Patterson function. *Acta Crystallographica Section A* **43**, 1-5.
- Thauer, R. K., Jungermann, K. & Decker, K. (1977) Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol Rev* **41**, 100-80.

- Thoden, J. B., Marti-Arbona, R., Raushel, F. M. & Holden, H. M. (2003) High-resolution X-ray structure of isoaspartyl dipeptidase from *Escherichia coli*. *Biochemistry* **42**, 4874-82.
- Thoden, J. B., Phillips, G. N., Jr., Neal, T. M., Raushel, F. M. & Holden, H. M. (2001) Molecular structure of dihydroorotase: a paradigm for catalysis through the use of a binuclear metal center. *Biochemistry* **40**, 6989-97.
- Turkenburg, J. P. & Dodson, E. J. (1996) Modern developments in molecular replacement. *Curr Opin Struct Biol* **6**, 604-10.
- Vagin, A. & Teplyakov, A. (1997) MOLREP: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography* **30**, 1022-5.
- Vagin, A. & Teplyakov, A. (2000) An approach to multi-copy search in molecular replacement. *Acta Cryst D* **56**, 1622-4.
- Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229**, 105-24.
- Vincent, F., Yates, D., Garman, E., Davies, G. J. & Brannigan, J. A. (2004) The three-dimensional structure of the *N*-acetylglucosamine-6-phosphate deacetylase, NagA, from *Bacillus subtilis*: a member of the urease superfamily. *J Biol Chem* **279**, 2809-16.
- Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I. & Joachimiak, A. (1999) MAD data collection - current trends. *Acta Crystallogr D Biol Crystallogr* **55**, 1726-32.
- Wang, B.-C. (1985) Resolution of phase ambiguity in macromolecular crystallography. In *Methods in Enzymology*, **115**, Elsevier Academic Press, 90-112.
- Wang, J., Hartling, J. A. & Flanagan, J. M. (1997) The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**, 447-56.
- Wang, Z. & Quirocho, F. A. (1998) Complexes of adenosine deaminase with two potent inhibitors: X-ray structures in four independent molecules at pH of maximum activity. *Biochemistry* **37**, 8314-24.
- Weiss, M. (2001) Global indicators of X-ray data quality. *Journal of Applied Crystallography* **34**, 130-5.
- Weiss, M. S. & Hilgenfeld, R. (1997) On the use of the merging R factor as a quality indicator for X-ray data. *Journal of Applied Crystallography* **30**, 203-5.

- Wendt, K. S., Schall, I., Huber, R., Buckel, W. & Jacob, U. (2003) Crystal structure of the carboxyltransferase subunit of the bacterial sodium ion pump glutaconyl-coenzyme A decarboxylase. *Embo J* **22**, 3493-502.
- Wilson, A. (1949) The probability distribution of X-ray intensities. *Acta Crystallographica* **2**, 318-21.
- Wilson, D. K., Rudolph, F. B. & Quioco, F. A. (1991) Atomic structure of adenosine deaminase complexed with a transition-state analog: understanding catalysis and immunodeficiency mutations. *Science* **252**, 1278-84.
- Yaffe, E., Fishelovitch, D., Wolfson, H. J., Halperin, D. & Nussinov, R. (2008) MolAxis: a server for identification of channels in macromolecules. *Nucleic Acids Res* **36**, W210-5.
- Zelder, O. (2008) Fermentative Herstellung von Chemikalien. *BIOspektrum* **04.08**, 348-50.
- Zhang, H., Yang, Z., Shen, Y. & Tong, L. (2003) Crystal structure of the carboxyltransferase domain of acetyl-coenzyme A carboxylase. *Science* **299**, 2064-7.
- Zhou, W., Bertsova, Y. V., Feng, B., Tsatsos, P., Verkhovskaya, M. L., Gennis, R. B., Bogachev, A. V. & Barquera, B. (1999) Sequencing and preliminary characterization of the Na<sup>+</sup>-translocating NADH:ubiquinone oxidoreductase from *Vibrio harveyi*. *Biochemistry* **38**, 16246-52.

## 7 Danksagung

Prof. Dr. Lars-Oliver Essen danke ich für die Möglichkeit, unter seiner kompetenten Anleitung in die Tiefen der Strukturbiologie vordringen zu dürfen. Besonders dankbar bin ich für das entgegengebrachte Vertrauen und die Freiheit, mich mit einer Vielzahl verschiedener Themengebiete befassen zu können.

Bei Prof. Dr. Wolfgang Buckel möchte ich mich für die sehr gute Zusammenarbeit auf dem Gebiet der Glutaconyl-CoA-Decarboxylase und die freundliche Übernahme des Zweitgutachtens herzlich bedanken.

Herrn PD Dr. Antonio Pierik danke ich für die Gelegenheit, im Rahmen des von ihm initiierten Enamidase-Projektes mit ihm zusammenarbeiten zu dürfen.

Daniela Bruegel, Dr. Ashraf Alhapel und Andrew Bozarth bin ich für die gute Zusammenarbeit bei den jeweiligen Kooperationsprojekten dankbar.

Außerdem möchte ich mich bei allen Kollegen aus der Arbeitsgruppe Essen für die freundschaftliche Arbeitsatmosphäre sehr herzlich bedanken.

Meinen Freunden danke ich für ihre moralische Unterstützung und die gemeinsam verbrachte Zeit abseits von Laboratorien und Synchrotron-Anlagen.

Ein ganz spezielles Dankeschön und einen dicken Schmatz meiner Liebsten Kathrin dafür, dass sie das alles bis zum Ende mit mir durchgestanden hat und sich nicht aus der Ruhe bringen lässt.

Bedanken möchte ich mich auch bei meiner Familie, besonders bei meinen Eltern Maria und Berthold, die all die Jahre an mich geglaubt und mich unterstützt haben.

Zu guter Letzt gilt meine Anerkennung Spike, der dann doch noch bis zum Schluss durchgehalten hat.